

PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR
DILEPTON MASS RESONANCES IN $H \rightarrow 4\ell$ DECAYS USING THE CMS DETECTOR AT
THE LHC

By

JAKE ROSENZWEIG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

© 2022 Jake Rosenzweig

This work is dedicated to the living and loving memory of Jacob Myhre.

ACKNOWLEDGEMENTS

Without so many two- and four-legged blessings along the way, I could never have made it to this point in my academic career. Thus, I begin by giving my endless gratitude to my high energy physics mentors, Professors Andrey Korytov and Guenkah Mitselmakher, for granting me this one-of-a-kind opportunity to do real *science* at CERN.

To my wife, Suzanne Rosenzweig, for showing me that dreams *do* come true. To my mother and father, Vicki and John, who always reassured me that I could achieve anything I put my mind to. Sleep peacefully, Mom. To my siblings, Alex, Ryan, Devin, Jace, and Claudia who frequently and gently reminded me that life existed outside of grad school. To Auntie Rachel and Uncle Yuri, who just *get* me and have given to me everything I could have possibly asked for. To my mentor, Sheldon Friedman, and his wife, Rita Friedman (Rosenzweig), who chose to invest in my success at a young age. I have only made it this far thanks to their undying encouragement, love, and optimism. To Sheldon's best friend, Dr. Bernard Khoury, whose reputation and has helped pave my own path. To the many moms who generously gave unconditional support during the darkest of times and unequivocal love during the brightest times: Cyndi Reilly-Rogers, Dawn Hood, Margaret Sherrill, and Silet Wiley.

To Dr. Filippo Errico for his focus, leadership, selflessness, and patience in leading the Higgs mass analysis. To Dr. Lucien Lo for showing me the simplicity and beauty of Python in his typical laid-back way and for leading the dilepton analysis. To Dr. Noah Steinberg for showing me how majestically physics can be communicated from mind to mind. To Dr. Darin Acosta, for spending many hours of physics discussion with me and the other students, who have helped us build our "CMS Office Hours" Ari Gonzalez, Cris Caballeros, Jeremiah Anglin, Sean Kent, Evan Koenig, Neha Rawal, Nik Menendez, John Rötter. To the gents who paved my way to and through the world of CMS, Brendan Regnery and Bhargav Joshi. To my mentee, Matthew Dittrich, for accepting the baton of knowledge and making everything come full circle.

To my comrades for showing me what it takes to survive the core courses, Dr. Atool Divakarla, Dr. Brien O'Brendan, Dr. Donyell Guerrero, and Dr. Vladinar Martinez. To Adamya Goyal for all his gentleness, humility, patience, and tutorage. To my Polish roommates in Saint-Genis-Pouilly for showing me what home away from home feels like: Bartoszek, Dziadziuś,

Karolina, and Sandruša. To the boys who have been there since the beginning: Jish, Willis, The Shane, Zacman, Duck, and Marcus for their clever competition and continual camaraderie which has shaped me to this day.

To Big Tree who stood as a symbol of strength, beauty, and life for centuries before us. As Irma's wild whirring winds worsened, the cacophony of ripping roots resounded throughout the western corridor. There I stood in that frozen moment—awestruck, speechless—watching her *fall* helplessly towards the physics building. What could have been a catastrophe of cataclysmic proportions was instead a gentle grazing against the north windows of NPB where, there, she was gracefully laid to rest.

Finally, to Existence, for this unpredictable, unbelievable blip of an experience called Life.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS	4
LIST OF TABLES.....	8
LIST OF FIGURES.....	11
ABSTRACT.....	12
CHAPTER	
1 INTRODUCTION	13
2 ALPHABET SOUP: H, SM, BEH, EWSB	17
2.1 Overview.....	18
2.2 Shortcomings of the SM	26
3 THE LARGE HADRON COLLIDER.....	28
3.1 Motivation	28
3.2 The LHC at CERN.....	28
3.3 The Journey of a Proton at the LHC.....	30
4 THE CMS DETECTOR.....	34
4.1 The Silicon Tracker	35
4.1.1 The Pixel Detector.....	36
4.1.2 The Strip Detector.....	36
4.2 The Calorimeters	37
4.2.1 Electromagnetic Calorimeter.....	37
4.2.2 Hadron Calorimeter.....	38
4.3 The Solenoid and the Steel Return Yoke.....	40
4.4 The Muon System	41
4.4.1 Cathode Strip Chambers	42
4.4.2 Drift-Tube Chambers	44
4.4.3 Resistive Plate Chambers.....	45
4.4.4 Gas Electron Multipliers	46
4.5 Trigger System	46
4.5.1 The Level-1 Trigger.....	46
4.5.2 The High-Level Trigger.....	46
4.6 Object Reconstruction.....	47
4.6.1 Electron and Photon Reconstruction	47
4.6.2 Muon Reconstruction	48
4.6.3 Tau Reconstruction	48
4.6.4 Jet Reconstruction	48
4.6.5 MET Reconstruction.....	48
5 HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL	60
5.1 Motivation	60

5.2	Analysis Overview	62
5.3	Analyzed Data	66
5.3.1	Triggers	67
5.3.2	Data Sets	67
5.3.3	Simulated Events	67
5.3.4	PileUp reweight	67
5.4	Event Selection	75
5.5	Background Estimation	77
5.5.1	Irreducible Background	77
5.5.2	Reducible Background	77
5.6	Techniques to Improve Mass Resolution	92
5.6.1	Event-by-event mass uncertainty	92
5.6.2	Model and procedure to derive corrections	98
5.7	Signal Modeling	102
5.7.1	Signal Normalization	102
5.7.2	Parameterizing the Signal Mass Line Shape	103
5.7.3	Building the 1D pdf	104
5.8	Systematic Uncertainties	105
5.9	Results	105
6	SEARCH FOR LOW-MASS DILEPTON RESONANCES IN THE $H \rightarrow 4\ell$ CHANNEL	106
6.1	Motivation	106
6.2	Overview	106
6.3	Data Sets, Simulated Samples, and Triggers	108
6.4	Event Selection and Reconstruction	108
6.5	Backgrounds	108
6.6	Systematic Uncertainties	108
6.7	Results	108
7	CONCLUSION	114
	REFERENCES	114
	BIOGRAPHICAL SKETCH	117

LIST OF TABLES

<u>Tables</u>	<u>page</u>
5-1 Trigger paths used to collect 2016 data (pre- and post-VFP)	67
5-2 Trigger paths used to collect 2017 data	68
5-3 Trigger paths used to collect 2018 data	68
5-4 Data sets used for 2016 UL.....	69
5-5 UL data sets used for 2017	70
5-6 UL data sets used for 2018	70
5-7 Names of simulated signal and background samples for 2016 data.	71
5-8 Names of simulated signal and background samples for 2017 data.	71
5-9 Names of simulated signal and background samples for 2018 data.	72
5-10 Yields in the narrow signal region $105 < m_{4\ell} < 140 \text{ GeV}$	75
5-11 Yields in the wide signal region $70 < m_{4\ell} < 170 \text{ GeV}$	76

LIST OF FIGURES

<u>Figures</u>	<u>page</u>
1-1 The elementary particles described by the SM.	13
1-2 Theoretical stability regions of the Universe.	15
2-1 Precision measurements of the W-boson mass performed by various collaborations.	18
2-2 The fundamental particles of the Standard Model.	20
2-3 A Feynman diagram showing electron-electron scattering, also known as Møller scattering.	21
3-1 (Left) A map image of the LHC ring. (Right) The CERN accelerator complex	29
3-2 Cross section of an LHC dipole magnet.....	32
3-3 Distribution of the average pile up during the LHC 2018 run.....	33
3-4 Diagram showing the constituents (<i>partons</i>) of two protons interacting in a pp collision... .	33
4-1 Life-size poster of the CMS detector.....	34
4-2 Points of interest along the LHC (Points 1–8).	49
4-3 Cut out of the CMS detector showing its various subdetector components.	49
4-4 A transverse view of CMS showing particle “filtration”.....	50
4-5 (Left) A simulation of the silicon tracker. (Right) The real silicon tracker within CMS....	51
4-6 A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved.	51
4-7 (Left) Cross-sectional view of the ECAL. (Right) An ECAL Dee.	52
4-8 (Left) ECAL crystals are grown in a lab. (Right) A single crystal.	52
4-9 Cross-sectional quadrant view of the HCAL components.	53
4-10 Cross section of CMS showing the magnetic field line strengths.	53
4-11 (Left) A simulated cut out view of the ME+ endcap, with the coordinate system of CMS in the top-left corner. (Right) The actual ME-2 disk of CMS is shown, revealing its ME-2/1 and ME-2/2 rings of CSCs.	54
4-12 Longitudinal cross section of CMS, showing the different pseudorapidity values (η) and also the different subdetector regions.	54
4-13 Exploded view of the cross section of a CSC showing how the 7 panels come together to form the 6 gas gaps between the panels.	55

4-14 (Left) A CSC with its top layer exposed. You can see very thin gold-plated tungsten wires which actually span the entire width of the CSC. Thicker vertical strips run along the length. (Right) More detail showing the radial strips and the horizontal wires. Also shown is the 5 segments of a CSC.	56
4-15 A muon passes through one of the gaseous layers of the CSC, ionizing the gas mixture and inducing a charge on the anode wires and cathode strips.	57
4-16 (Left) Comparators are used to compare neighboring strip cluster charge to determine on which half-strip the peak charge resided. (Right) A muon passes through all six layers of a CSC inducing charge on various half-strips.	57
4-17 A drift-tube chamber with 3 superlayers (SLs) is sandwiching a honeycomb plate. Two SLs are indicated by the blue rectangles, whose cells (orange rectangle) have anode wires (red dots) oriented parallel to the beam pipe. The third SL is indicated by the green rectangle and is staggered orthogonally to the other two SLs.	58
4-18 A cross section of a DT cell showing the drift lines (light blue), isochrone lines (dark blue), dimensions of the cell, and locations of the anode wire and cathode strips.	58
4-19 A cross section of CMS showing the locations and numbering scheme of the drift tubes in the barrel.	59
5-1 Various measurements of m_H made by the CMS and ATLAS collaborations.	61
5-2 Branching ratios of Higgs boson decays as a function of the Higgs boson mass.	63
5-3 Feynman diagrams for the $gg \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$ and $gg \rightarrow ZZ^* \rightarrow 4\ell$ processes.	64
5-4 Distribution of $m_{4\ell}$ from $H \rightarrow ZZ^* \rightarrow 4\ell$ events using Full Run 2 data.	65
5-5 words	74
5-6 Distribution of the four-lepton invariant mass for each final state.	76
5-7 Diagram showing how RB prompt-lepton processes contribute to observed PTS-lepton events.	81
5-8 Distributions of the predicted and observed RB yields for the 2P2F CR (2016 pre-VFP)... .	83
5-9 Distributions of the predicted and observed RB yields for the 2P2F CR (2016 post-VFP). .	84
5-10 Distributions of the predicted and observed RB yields for the 2P2F CR (2017).....	85
5-11 Distributions of the predicted and observed RB yields for the 2P2F CR (2018).....	86
5-12 Distributions of the predicted and observed RB yields for the 3P1F CR (2016 pre-VFP)... .	87
5-13 Distributions of the predicted and observed RB yields for the 3P1F CR (2016 post-VFP). .	88
5-14 Distributions of the predicted and observed RB yields for the 3P1F CR (2017).....	89

5-15 Distributions of the predicted and observed RB yields for the 3P1F CR (2018).....	90
5-16 Electron misidentification rates vs. electron p_T for the OS Method.	93
5-17 Muon misidentification rates vs. muon p_T for the OS Method.	94
5-18 TODO	95
5-19 words.....	96
5-20 words.....	97
5-21 words.....	98
5-22 TODO	100
5-23 TODO	101
5-24 TODO	102
5-25 TODO	103
5-26 TODO	105
6-1 Exotic decays of the Higgs boson to the 4ℓ final state.....	107
6-2 test.....	110

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR
DILEPTON MASS RESONANCES IN $H \rightarrow 4\ell$ DECAYS USING THE CMS DETECTOR AT
THE LHC

By

Jake Rosenzweig

December 2022

Chair: Andrey Korytov
Co-Chair: Guenakh Mitselmakher
Major: Physics

The mass of the Higgs boson is measured in the $H \rightarrow ZZ^* \rightarrow 4\ell$ ($\ell = e, \mu$) decay channel and is found to be $m_H = 125.38 \pm 0.11$ GeV; the most precise measurement of m_H in the world to date. The data for the measurement were produced from proton-proton (pp) collisions at the Large Hadron Collider with a center-of-mass energy of 13 TeV during Run 2 (2016–2018), corresponding to an integrated luminosity of 137.1fb^{-1} , and were collected by the Compact Muon Solenoid experiment. This measurement uses an improved analysis technique in which the final state muon tracks are constrained to originate from the primary pp vertex. Using data sets from the same run, a search for low-mass dilepton resonances in Higgs boson decays to the 4ℓ final state is also conducted. No significant deviation from the Standard Model prediction is observed.

CHAPTER 1 INTRODUCTION

The Universe, while overwhelmingly vast, is built from a remarkably few kinds of elementary (i.e., indivisible) particles. As shown in Fig. 1-1, any elementary particle can be classified into 1 of these 3 categories: matter particles (*fermions*), force-carrying particles (*gauge bosons*), and Higgs bosons. There are 12 kinds of fermions which can be split evenly into 2 groups, depending on with which forces they interact: those that interact via the electromagnetic (EM) force and the weak nuclear force are classified as *leptons*, of which there are 6 kinds (“*flavors*”), whereas those that interact via the EM, weak, *and* strong nuclear forces are classified as *quarks*, of which there are also 6 flavors. There are 4 kinds of gauge bosons, each of which is a force carrier for a specific force (the gluon is said to *mediate* the strong force, the photon mediates the EM force, while both the W^\pm and Z bosons mediate the weak force). Thus, all the diversity and manifestations of reality come from only 17 kinds of “building blocks”. It is the mission of particle physicists to understand the underlying mathematical structure that describes Nature in as accurate—and concise—a theory as possible. The best theory that stands today is called the Standard Model (SM) of particle physics.

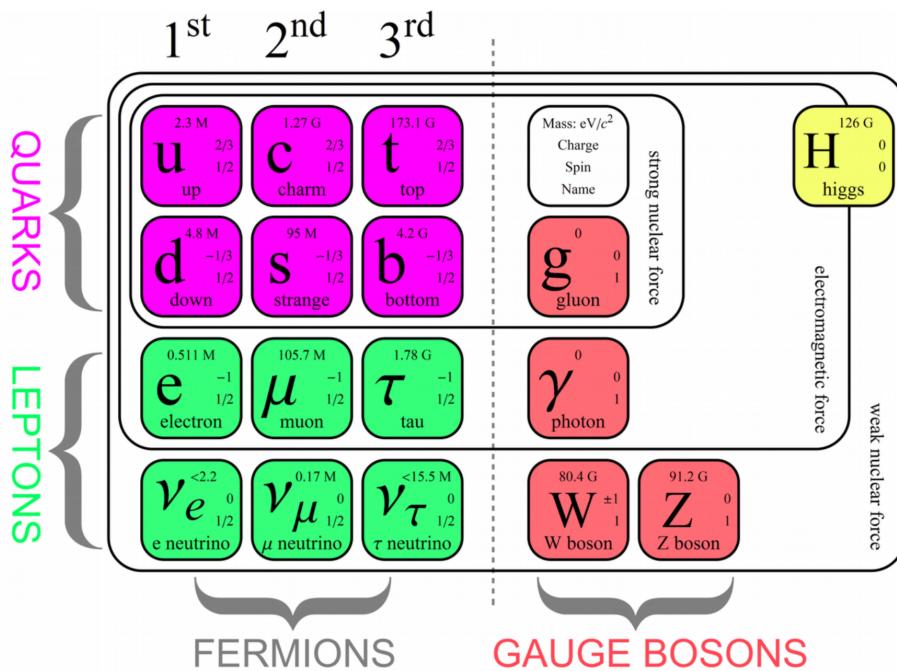


Figure 1-1. The elementary particles described by the SM.

The SM has bore witness to many triumphs over its approximately 70-year development: it predicted the existence of quarks which were experimentally confirmed in the mid-1970s; it predicted the existence of the tau neutrino which was experimentally confirmed in 2000; its most groundbreaking prediction was experimentally confirmed on July 4, 2012—just shy of 10.5 years ago from the date of this dissertation writing—when the ATLAS and CMS collaborations announced the discovery of the Higgs boson [1–3].

Quantum Field Theory is the mathematical and conceptual backbone of the SM. Within its framework, all particles are *excitations* of their corresponding field so, e.g., *every* electron in the Universe is thought to be an excitation of the single electron field that permeates all of spacetime. The existence of the Higgs boson (H) suggests that its corresponding field exists—the *Higgs field*—and, thus, H is the quantum excitation of that field. This all-pervasive Higgs field and its corresponding boson are generated mathematically via the Brout-Englert-Higgs (BEH) mechanism. Most SM particles—except for neutrinos, photons, and gluons—“acquire” their mass by interacting with the Higgs field. This is also how the Higgs boson itself acquires its mass (m_H); by interacting with its own field!

The masses of SM particles *depend* on the value of m_H . . . so what is its value? Unfortunately, m_H is a free parameter of the SM, so theory is unable to provide a value for m_H based solely on other fundamental constants. Instead, the value of m_H must be measured by experiment—and has been measured multiple times [4–6]. Although m_H has already been measured, it is important to continually improve the measurement by lowering the uncertainties on (i.e., increasing the precision on) the mass value. A more precise value of m_H is two-fold: first, it improves the theoretical limits on the masses of the other elementary particles and, second, it very well may determine the stability of our universe, as shown in Fig. 1-2.

The production of a Higgs boson is only feasible at conditions close to those thought to exist at the beginning of the Universe. This grand achievement is accomplished frequently by the Large Hadron Collider (LHC), located on the border of France and Switzerland. It is the largest and most powerful proton-proton (pp) collider ever made. The LHC accelerates protons to incredible

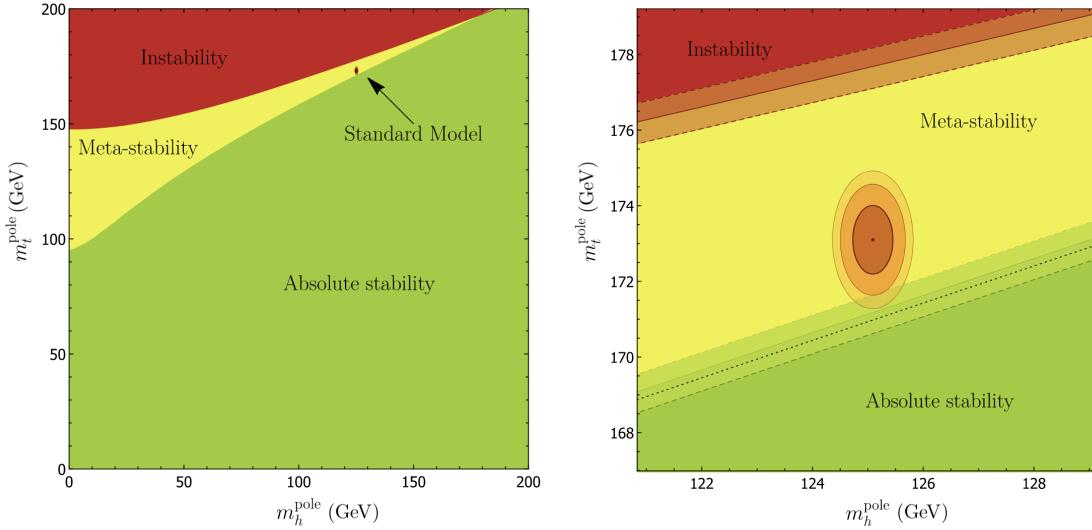


Figure 1-2. (Left) Theoretical stability regions of the Universe based on the pole masses of the top quark (m_t^{pole}) and Higgs boson (m_h^{pole}). (Right) A closeup of the SM region of the left plot. The contours represent the 68%, 95%, and 99% confidence levels based on the experimental uncertainties of m_t^{pole} and m_h^{pole} . Plots taken from [7] and units added to all axes.

speeds, very close to the speed of light. When these fast-moving protons collide, the pp collisions can have a center-of-mass energy as high as 13 TeV. Newly produced particles spew out of the collision points and are analyzed by detectors like the aforementioned ATLAS and CMS experiments. These enormous particle detectors have thousands of scientists performing dozens of analyses to look for hints of beyond Standard Model (BSM) physics, extra dimensions, miniature black holes, and more.

This dissertation utilizes data collected by the CMS experiment during the LHC Run 2 (2016–2018) to perform the world’s best precision measurement of m_H to date. This new measurement utilizes the following improvements compared to previous measurements:

- Nearly four times as much collected data from Run 2 ($L_{\text{int}} = 137.1 \text{ fb}^{-1}$) vs. the data used for the 2016 measurement ($L_{\text{int}} = 35.9 \text{ fb}^{-1}$).
- Four final-state categories: 4 μ , 4e, 2e2 μ , 2 μ 2e. In previous measurements, the last two final states (the mixed-flavor states) were combined, when truly they have different kinematical properties (depending on into which lepton pair the Z_1 decayed): different peak widths (instrumental resolutions), different signal efficiencies, and different relative levels of

reducible background.

- Ultra-Legacy (UL) reconstruction for muon, electron, photon, and jet tracks. This significantly improves electron momenta and improves the other particle momenta, though to a lesser degree.
- The measurements of muon p_T are improved by constraining the muon tracks to originate from the interaction vertex (also called a *vertex constraint*).
- When extracting the value of m_H in past measurements, a 3D pdf $(m_{4\ell}, \sigma_{m_{4\ell}}, \mathcal{D}_{\text{bkg}}^{\text{kin}})$ was built into a factorized form $f(m_{4\ell}, \sigma_{m_{4\ell}} | m_H) \cdot g(\mathcal{D}_{\text{bkg}}^{\text{kin}} | m_{4\ell})$, which was later found to contain an existing correlation between $\sigma_{m_{4\ell}}$ and $\mathcal{D}_{\text{bkg}}^{\text{kin}}$. To account for this correlation, now the events are split into 9 categories based on the per-event *relative* mass uncertainty $\left(\frac{\sigma_{m_{4\ell}}}{m_{4\ell}}\right)$ and, for each, a 2D pdf $(m_{4\ell}, \mathcal{D}_{\text{bkg}}^{\text{kin}} | m_H)$ is built.
- The systematic uncertainties on electron and muon momentum scales ($p_T^{e,\mu}$) are reduced, thanks to a more detailed analysis on the uncertainties. This has the additional effect of significantly reducing the uncertainty on the per-event four-lepton mass resolution.

The following chapters of this dissertation begin by describing the function and engineering of the Large Hadron Collider in Chapter 3. Then, a thorough description of the CMS Experiment and its composite subdetectors is given in Chapter 4. Next, the details of the precision measurement of the Higgs boson mass using the LHC Run 2 data is discussed in Chapter 5. Finally, the results of the Higgs boson mass measurement analysis is summarized in Chapter 7.

CHAPTER 2

ALPHABET SOUP: H, SM, BEH, EWSB

The Standard Model (SM) is a collection of the most accurate and self-consistent particle physics theories that mathematically describe the properties of particles within the universe and their interactions with each other. For the past century, some of the most brilliant minds in physics have spent their entire careers to develop equations, mathematical tricks, and completely novel ideas to help build a solid foundation for the SM. To demonstrate the unparalleled accuracy with which the SM predicts physical phenomena, one can compare the experimentally measured anomalous magnetic moment of the electron

$$a_e^{\text{exp}} = 0.001\ 159\ 652\ 180\ 73(28)$$

to the value predicted by the SM

$$a_e^{\text{pred}} = 0.001\ 159\ 652\ 181\ 643(764).$$

An impressive agreement to better than one part per trillion.

On the other hand, the SM has no explanation for some observed physical phenomena (Sec. 2.2), e.g., the existence of dark matter [8], and is thus not a fully descriptive model of the Universe. Furthermore, the Collider Detector at Fermilab (CDF) Collaboration recently measured the mass of the W boson to be $80433.5 \pm 9.4 \text{ MeV}$, whereas the SM expectation $80357 \pm 4 \text{ MeV}$ —a discrepancy of 7.0 standard deviations [9]—as is shown in Fig. 2-1. Recent results from the Baksan Experiment on Sterile Transitions suggest electron neutrinos oscillating between sterile neutrino states [10]. Nevertheless, the SM has laid important foundations in particle physics for over 100 years and is fully explained with the discovery of the Higgs boson in 2012 TODO:CITE.

The remainder of this chapter presents a general overview (Sec. 2.1) of the particles and forces described by the SM, followed by a brief analysis of the mathematical underpinnings of the SM (Sec. ??), and then finally a brief description of the shortcomings of the SM (Sec. 2.2).

SM OF PARTICLE PHYSICS CHAPTER OUTLINE OVERVIEW: - Main concept: -
Everything is particles! Fields permeate all of space. Their mathematical treatment is handled in Sec.TODO. Particles are excitations of the fields. Interactions are between particles Bosons

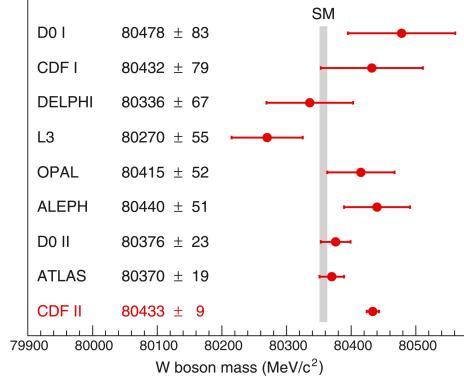


Figure 2-1. Precision measurements of the W-boson mass performed by various collaborations.

Fermions Leptons Quarks MATHEMATICAL FRAMEWORK - Show SM Lagrangian. - Derive interactions between particles. BEH MECHANISM - Robert Brout, François Englert, and Peter Higgs claimed that elementary particles acquire their masses via spontaneous EWSB. (reword!) - Sheldon Glashow, Abdus Salam, and Steven Weinberg were able to unify the weak and electromagnetic forces into a single force, above the unification energy of 246 GeV. - In 1973 the Gargamelle collaboration experimentally confirmed the existence of the electroweak force by discovering neutral currents in neutrino scattering experiments. - Furthermore, in 1983 the UA1 and UA2 collaborations used proton-antiproton collisions to discover the theorized W and Z electroweak gauge bosons. SHORTCOMINGS OF THE SM - Each of these is a new paragraph?: - Neutrino oscillations confirm that neutrinos *do* have masses but interaction with the Higgs field is not responsible for this origination. - No gravity. - Why matter and no antimatter? - Where does dark matter fit in? - Why should there be exactly 3 generations of fermions? - Although the SM does not answer any of the above questions,

2.1 Overview

The SM is a non-Abelian set of gauge theories. Non-Abelian gauge symmetry gives rise to charged gauge bosons.

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

Charge and hypercharge both obey U(1) symmetry.

Weak hypercharge, weak isospin (L reminds us that this interaction only effects left-handed fermions), color charge Q : *What is the difference between $U(1)_q$ and $U(1)_Y$?* Q : *Does L stand for left-handed?*

All observers must Lagrangian The connection between conservation laws and symmetries is best analyzed via Lagrangian mechanics.

Each Lagrangian has its own set of “Feynman rules”.

The Particles: The Players on the Fields

Contrary to intuition, fundamental particles are not hard, billiard-ball-like objects as is often perceived. Instead the SM predicts that every particle is actually an *excitation* of its corresponding field. So for example the electron is an excitation of the *electron field*, $\psi_e(x)$, a bispinor that describes a spin-1/2 field which follows the Dirac equation:

$$(i\hbar c \gamma^\mu \partial_\mu - mc^2)\psi(x) = 0.$$

In fact, *every* electron is an excitation of this same electron field. An electron is identical to every other electron in every way (same mass, same charge, etc.). Quantum Mechanics (QM) does a fine job of describing how slow-moving particles behave. However, as soon as particles begin moving at 30.5% of the speed of light (91.4 E6 m/s), then there is a 5% difference between the particle’s rest mass energy ($E = mc^2$) and its total relativistic energy ($E = \gamma mc^2$). In other words, fast-moving particles must account for effects due to Special Relativity (SR). The merger of QM and SR gives rise to Quantum Field Theory (QFT) - the backbone of the SM.

Fig. 2-2 shows all the fundamental particles that have been discovered up to the present day. The phrase “fundamental particle” just means that the particle is not composed of anything smaller than itself. These particles are not just diabolical creations from theorists. No, these particles are precisely defined, mathematical objects whose existence has been predicted by the SM and experimentally verified time and time again in the laboratory.

Each particle has a unique set of properties (like mass, electric charge, spin, etc..) that distinguish it from all the other particles. It is one of the primary goals of particle physics to

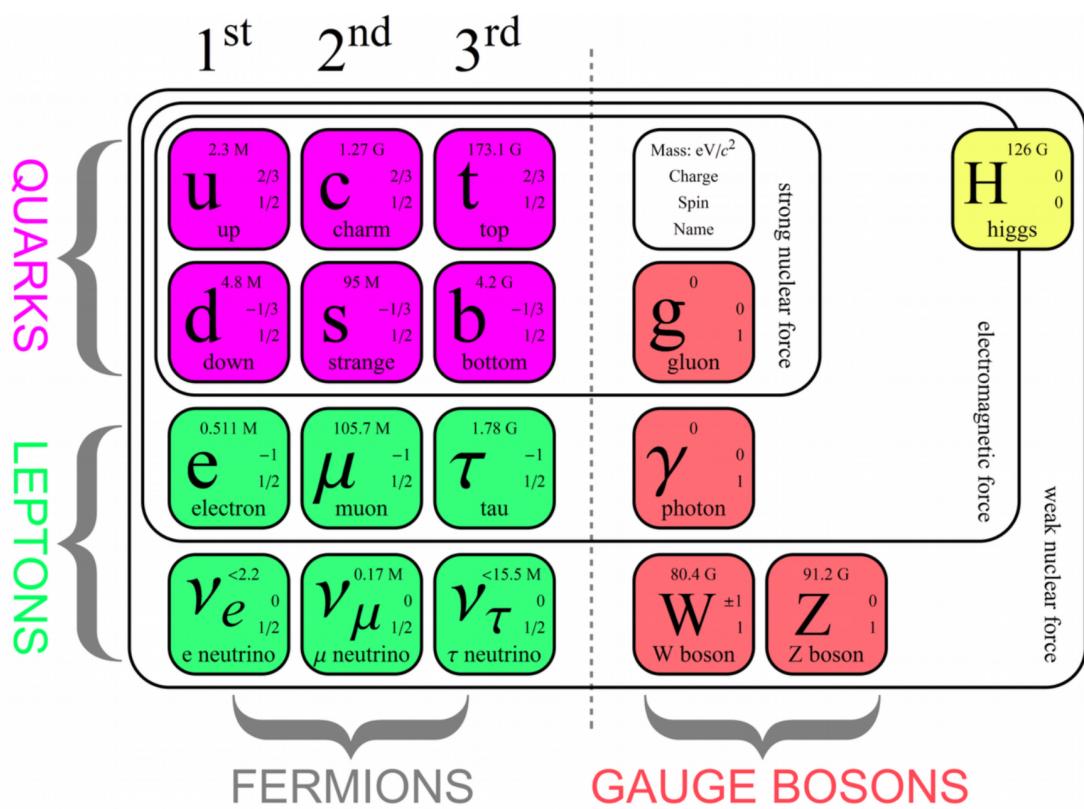


Figure 2-2. The fundamental particles of the Standard Model.

determine these properties, because ultimately their properties determine their *interactions* with one another.

Without further ado, let's meet the particles. There are two major types: *bosons* and *fermions*. We are going to take a non-traditional route and introduce the bosons first, then the fermions.

Bosons: Use the Force

Ever wonder how two electrons "know" that they are near each other and that they should repel? Fig. 2-3 shows a Feynman diagram of two electrons "communicating" with each other by means of an intermediate photon. I like to think of it as the two electrons "playing catch" with the photon. The first electron recoils from the throw and then the second electron recoils from catching the photon. The photon carries some momentum away from the first electron and brings it to the second one, therefore making it look as if the two electrons are repelling one another! The photon isn't *real* of course - it is said to be a *virtual* photon. Now we see why bosons are called the force carriers.

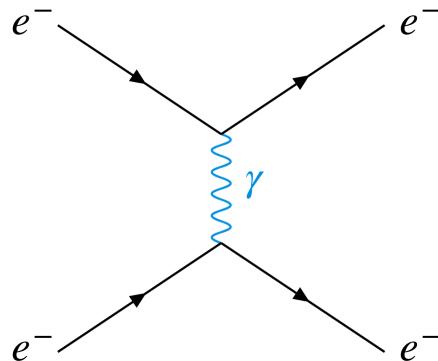


Figure 2-3. A Feynman diagram showing electron-electron scattering, also known as Møller scattering.

A diagram such as the one above is a Feynman diagram and it gives us a wonderfully simple way to visualize particle physics processes. It's not *actually* what happens between the particles, but it is a good starting point. Each diagram is actually a single, scalar number - a complicated QFT integral that tells you how likely a process is to happen. Another benefit to Feynman Diagrams is that they are kind of like tinker toys, in that you can string them together in novel

ways to predict real-world processes. Quantum Electrodynamics (QED), one of the theories that make up the SM, mathematically describes and predicts the electrons mediating such a photon between them. It's not just limited to electrons mediating photons, however. QED and other QFTs can predict to astounding accuracy how likely a process is to happen, between whichever particles and fields. You just need to know their properties first.

We have now met the first force carrier: the photon. It is a massless particle and is the mediator of the EM force. Photons only interact with particles that carry *electric charge*. Depending on what kind of charge a particle carries, determines with which bosons it may interact and via which forces. Speaking of forces, the four fundamental forces found in nature, along with their decreasing, relative strength are:

1. strong force (1)
2. EM force (10^{-1})
3. weak force (10^{-6})
4. gravitational force (10^{-40})

If the photon is the mediator of the EM force, then what mediates the other forces?

The mediators of the strong force are the **8 gluons**. Similar to the photon, they are also massless, but that's about all they have in common. Gluons are trapped inside of protons, neutrons, and other hadronic matter. They are responsible for “glueing” nuclei to together, hence their name. Just as photons can only interact with particles that have electric charge, gluons can only interact with particles that have *color* charge. Interestingly, gluons themselves carry color charge which they mediate back and forth between quarks (fermions discussed below). This is quite different from the photon which itself does not carry electric charge. There are three kinds of color charges: red, green, and blue. Every gluon carries two color charges: one kind of color and an anticolor: antired, antigreen, or antiblue.

There are three bosons which mediate the weak force: the Z, W^+ , and W^- . They are extraordinarily massive particles, weighing in at 91.2 GeV for the Z and 80.4 GeV for both kinds

of W bosons. That means the W bosons weigh more than an iron atom! These bosons interact with any particle that carries “weak hypercharge”. The weak force has plagued physicists for nearly a century until only recently. Particles which decay via the weak force live an astonishingly long time. Take for example the neutral pion (π^0). It decays very quickly, via the EM force, into two photons on the order of 10^{-18} s. Now take the charged Kaon K^+ . This particle decays into three charged pions, but takes on average 10^{-8} s to do so. Over 10 orders of magnitude different from the pion decay. This is because the charged kaon decays via the weak force.

The last boson not yet mentioned is the scalar Higgs boson, which is introduced in Sec.TODO below.

Fermions: Each One Matters There are 12 kinds of fermions - the matter particles of the Universe. They comprise all the “stuff” that we see and feel. All fermions have half-integer spin, typically a value of 1/2. The fermions can be split into two groups depending on if they interact with the strong force (quarks) or not (leptons). Let’s consider the leptons first.

Leptons: We already introduced one lepton earlier: the electron. Looking again at the “particular table”, the electron has a heavier brother, the muon, which is 200 times heavier than the electron. Then there is an even heavier sibling: the tauon. All three of these leptons have the familiar -1 charge which allows them to interact via the EM force and exchange photons with other electrically charged particles.

The charged leptons also carry weak hypercharge, which allows them to interact via the weak force. If a charged lepton interacts with a W^\pm boson, it can transform into its corresponding “partner” - the other member of the $SU(2)$ isospin doublet: the neutrino. These fickle particles are neutral and *only* interact via the weak force (well, and maybe gravity). They are very difficult to detect.

Quarks: The six quarks are the fermions which interact with gluons. They have *quarky* names like: up, down, charm, strange, top, bottom. These are called the six “flavors” of quarks. The top quark is an absolutely massive particle, reaching the top of the mass scale of any particle at 173 GeV - about as heavy as a tungsten atom.

Quarks are electrically charged particles, but they have fractional charge. Each quark in the top row of Fig. 2-2 has +2/3 electric charge and the bottom row has -1/3. That's why when you combine two up quarks with a down to form a proton quark, the combination of electric charge yields +1.

Just as the leptons carried weak hypercharge and could interact via the weak force, so too can quarks. The W^\pm bosons can change one flavor of quark into another. The Z boson only affects the spin, momentum, and energy of the particle with which it interacts.

In addition to electric charge, quarks also carry one kind of color charge, either red, green, or blue. It is this color charge which allows them to interact with gluons via the strong force. This is an artifact of being gauge bosons of the $SU(3)$ symmetry group. They combine in different ways to form at least two types of hadrons. The first type is baryons, like protons, neutrons, lambdas (anything that is qqq) and the second type is mesons, like pions, kaons, etas (anything of some form like $q\bar{q}$). For some reason which is not completely understood, only colorless bound states form in nature. Just as a '+' charge would negate a '-' charge, so too would the 'red' color charge negate 'antired' (as in the case of an observable meson) or even combining red, green, and blue (as in the case of a baryon) would yield a colorless bound state.

Antiparticles: It should be noted that almost every *particle* has a corresponding *antiparticle*, whose charges (e.g., color charge, electric charge) are all opposite the original particle's charges. Accounting for leptons, quarks, bosons, bound states of quarks, and now antiparticles, it is easy to see why sometimes particle physics is referred to as a "zoo"!

Electroweak Symmetry Breaking and the Higgs Boson At large energy scales, like those found during the Big Bang or those produced in the energetic proton-proton collisions of the Large Hadron Collider, the electromagnetic (EM) force and weak nuclear force are one and the same: they are unified. However, at lower temperatures (like room temperature for example), we know that the weak force is very different from the EM force. The former mediates decays of radioactive substances, whereas the latter mediates the excitation of electrons in an atom. So what is responsible for the separation of these two forces, this so-called *electroweak symmetry breaking*?

Upon writing down the equations of motion from the SM Lagrangian (easier said than done), one discovers that all the particles mentioned earlier should have *no mass*. Well that's a problem because most particles in nature definitely have mass, like the quarks, leptons, W^\pm , and Z bosons. By introducing a complex SU(2) doublet of scalar fields into the SM Lagrangian, in such a way that it leaves the Lagrangian invariant, then all peace can be restored. This scalar field turns out to be the Higgs field, and its excitations are Higgs bosons. Doing so reveals the particle which should have mass. The process of introducing a Higgs field and breaking the electroweak symmetry is called the **Higgs Mechanism**.

Each particle interacts with the Higgs field with a different strength: in fact, a particle's coupling strength to the Higgs field is exactly its mass! The more the particle interacts with the Higgs field, the more mass it gains. Excitations, or quanta, of the Higgs field are Higgs bosons and are a direct consequence of introducing a Higgs scalar field into the SM Lagrangian to allow particles to have mass.

- strong force (1)
- EM force (10^{-1})
- weak force (10^{-6})
- gravitational force (10^{-40})

Speaking of forces, the four fundamental forces found in nature, along with their decreasing, relative strength are:

Electroweak Interaction Particles can interact with one another at long range. For example, an electron can emit a photon which can travel a

Electromagnetic force and weak force were unified Electroweak symmetry. This introduced 4 electroweak bosons: the photon, Z , W^+ , and W^- . Then lectroweak symmetry breaking happened. This made the photon massless and the Z , W^+ , and W^- as (very) massive.

Higgs Mechanism Force field that fills space in the whole Universe but has no source or direction. The field has the same field at every point. It's a scalar field with spin 0. It does have electroweak charge.

2.2 Shortcomings of the SM

The SM has only mathematically accommodated the strong, EM, and weak forces. One problem however is that the SM can't predict the mass of the Higgs boson... or the mass of *any* particle for that matter. That's not the only thing the SM has trouble doing. For example, the SM can't...

- ...incorporate gravity into its mathematical framework.
- ...explain why most of the Universe is made of matter and very little antimatter.
- ...predict the existence of dark matter - but we know it's there from observation.
- ...explain why there should be exactly three generations of fermions.

No Gravity. Can't combine quantum mechanics and gravity. No neutrino masses. Are they Dirac or Majorana particles? Higgs field parameters appear highly fine-tuned. Does not explain dark matter or dark energy.

So we see that the SM isn't the ultimate Theory of Everything, but it does a pretty good job. How can we test the SM and try to break it or confirm it? There are at least two routes to choose from: A patient route and an impatient route. The patient route requires us to wait until our particles of interest maybe come from outer space or, if we produce it in the lab, wait for it to decay into other particles. This could take a VERY long time, (possibly way longer than the age of the Universe - if a proton even decays at all!), or it could take as short as a billionth of a billionth of a millionth of a second, like in the case of a Z boson. It's not the most reliable method, it is difficult to control, and it requires a lot of patience. - Origin of mass: yes, the Higgs boson—but is it certainly the same Higgs boson as predicted by the SM? - SUSY: - A way to unify the fundamental forces (all 4?) but SM doesn't predict SUSY particles. - Do they exist? No evidence yet.

- Dark matter/Dark Energy: - Rotational velocity data from the outer reaches of our own galaxy suggest that there is much more matter in the universe than what has been directly observed. The SM has no explanation for dark matter/dark energy.

- Matter/Antimatter Asymmetry: - The Big Bang supposedly (TODO) produced equal amounts of matter and antimatter. - So why is the universe made of *only* matter?

Instead, let's be impatient: let's smash particles together and convert their energies into new kinds of matter. If we use hadrons, which are made of smaller parts like, quarks and gluons (let's call them "partons") then we will have many more interactions and a lot more fun (Fig. 3-4). We are going to need a lot of energy, so we should make a large collider. Let's make a Large Hadron Collider!

CHAPTER 3

THE LARGE HADRON COLLIDER

3.1 Motivation

Although the SM has shown to be an astoundingly accurate framework so far, it must continue to be scrutinized by the barrage of measurements that either confirm or contradict its predictions. Interestingly, a recent measurement of the mass of the W boson has shown significant deviation from SM predictions, with a sensitivity of 7σ [9]. After all, undeniable fact comes from the reproducible results obtained from *measurement*—not from some theoretical model which *may* or *may not* describe reality. Whenever the predictions of a model directly contradict the results from measurement, the model must necessarily be cast aside and replaced by one whose predictions concur with the results of measurement.

So how *are* measurements obtained in the realm of particle physics? Modern day physicists study the fundamental constituents of matter and their interactions by using state-of-the-art technologies combined with time-tested methodologies: by smashing tiny bits of matter together to turn them into even *tinier* bits. Such is the purpose of the world’s largest and most powerful particle accelerator—the Large Hadron Collider (LHC).

3.2 The LHC at CERN

Deep beneath the surface of the earth (50–175 m), the LHC straddles the border shared by France and Switzerland. Sandwiched between the scenic Jura mountains to the northwest and the sprawling city of Geneva (French: Genève) to the southeast is CERN. To illustrate the enormous circumference (26.659 Km) the sheer size of this circular accelerator, Fig. 3-1 (Left) shows the LHC drawn on a map. For reference, the inscribed area of the LHC (56.7 Km^2) is almost four times greater than the area of the neighboring city of Geneva (15.9 Km^2).

The LHC is not only a particle *accelerator* but also a proton-proton (pp), proton-lead ion, and lead-lead ion *collider*. By sending one particle beam clockwise around the ring and the other beam counterclockwise, the charged particles are carefully maneuvered around the ring using dipole magnets and collimated into tight proton *bunches* using quadrupole magnets before they ultimately collide at 4 specific points along the LHC, as shown in Fig. 3-1 (Left, red stars). When the LHC is fully powered, *each* proton in the beam carries an average energy of 6.5 TeV which

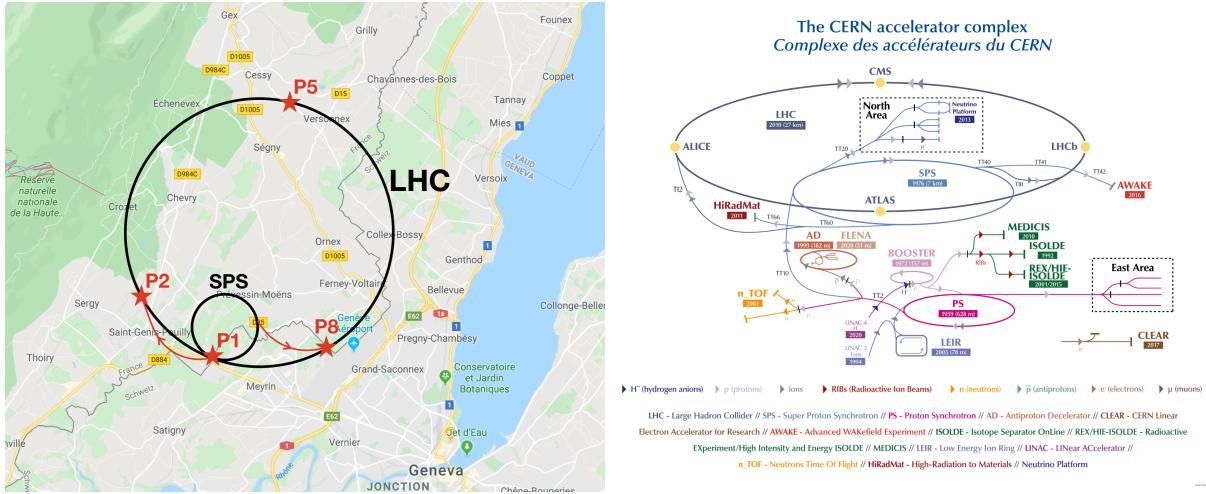


Figure 3-1. (Left) The LHC ring (bigger ring) and the Super Proton Synchrotron (smaller ring) with the nearby town of Geneva for size comparison. The four red stars indicate the pp collision points. (Right) The accelerator complex at CERN.

gives a single pp collision a center-of-mass energy of 13 TeV. This emulates the conditions theorized to exist at the beginning of the universe, which allows cosmological studies to be carried out. The hugely energetic pp collisions cause the quark and gluon constituents within the protons to interact with each other and transform into new particles. The newly created particles and the residual particle debris are ejected away from the collision point—whether straight down the beampipe, completely orthogonal to it, or somewhere in between. A massive particle detector is stationed at each of the 4 collision points to detect the outgoing particle “spray”. The 4 main particle detectors and their locations along the LHC are:

- A Toroidal LHC ApparatuS (ATLAS)—located at the first collision point (P1),
- A Large Ion Collider Experiment (ALICE)—located at P2,
- the Compact Muon Solenoid (CMS, Chapter 4) experiment—located at P5, and
- the LHC-beauty (LHCb) experiment—located at P8.

The world-renowned feat of digging the tunnel for, constructing, commissioning, and monitoring the LHC was made possible by CERN: the European Organization for Nuclear Research (French: *Conseil Européen pour la Recherche Nucléaire*). CERN is an international collaboration of—at the time of this writing—more than 33 countries, each of which is considered either a Member

State, an Associate Member State, or an Observer. The complex of CERN (Fig. 3-1, Right) is located just to the west of P1 and is akin to a small science *city*—complete with many offices, manufacturing facilities, and experiments such as the Antiproton Decelerator (AD), the Neutrons Time of Flight (n_TOF), and the Isotope Separator OnLine (ISOLDE) experiments. Although the LHC is the most famous of the accelerators at CERN, its fame is only made possible by a series of smaller and lesser-known accelerators that *feed* the LHC. Therefore, a natural way to explore the intricacies and inner workings of the LHC is to follow the path of one of its “inhabitants”—a single proton—as it makes its way to and through the gigantic collider.

3.3 The Journey of a Proton at the LHC

The journey to discovery begins in a surprisingly small tank of hydrogen gas (H_2) located in the LINAC4 building at the main CERN site. Inside this tank, a proton—conveniently called P. Roton—and approximately 10^{23} other protons coexist in bound states as molecules of H_2 . Although the tank has a meager mass of 10 Kg, it has enough protons inside to keep the LHC colliding protons for over *200 000 years* of constant operation.

Protons get injected into a series of increasingly larger accelerators. It begins with the *Linac4*—a linear particle accelerator—that accelerates hydride ions (H^-) to 160 MeV which eventually make their way into the *Proton Synchrotron Booster* (PSB). In the PSB, each hydride ion has its electron pair completely stripped away, leaving only the bare proton. The protons then enter a series of circular accelerators, each machine feeding protons into the next while increasing the proton center-of-mass energy by at least 1 order of magnitude.

- Protons are then accelerated to 2 GeV at which point they are injected into the *Proton Synchrotron* (PS).
- The PS then increases the proton energy to 26 GeV to be fed into the *Super Proton Synchrotron* (SPS).
- The penultimate step is for the SPS to further energize the protons to a center-of-mass energy of 450 GeV.
- Finally the protons enter the LHC and are further accelerated to the maximum energy of 6.5 TeV using RF cavities to kick them.

The protons would travel in a straight line if not for the 1232 dipole magnets sprinkled along the LHC. They are made out of copper-clad niobium-titanium, require the cryogenics of 96t of superfluid helium-4 to ultimately turn the proton beams. There are also 506 quadrupole magnets that compress the proton bunches to make them as linear as possible before they collide. Finally two incoming proton bunches approach a common collision point. Beam pipes are conjoined in an “X” shape where 2 proton bunches cross—a *bunch crossing* (BX). Out of more than 40 million pp collisions that could have occurred, a mere 50 collisions take place on average (i.e., only 0.0001%). This is a testament to just how small protons truly are.

Just as the PS feeds protons into the SPS, which feeds protons into the LHC, so too is it being considered for the LHC to feed a new project—the 100 Km Future Circular Collider.

In the event that P. Roton does not collide with any of the oncoming protons, then it is simply “recycled” and continues going around the LHC ring for another opportunity at a pp collision.

Contrary to what some people may think, protons are not sent one by one at each other, hoping for a collision. Instead 100 billion protons are packed together into a “proton bunch”. A single proton bunch is about the size of a human hair ($\approx 50\mu\text{m}$ wide and $\approx 10\text{cm}$ long). The clockwise and counterclockwise rings are filled to a maximum of 2808 proton bunches, each one spaced 25 ns apart, and then sent to collide.

It requires an incredibly strong magnetic field to turn the protons as they make their revolutions around the LHC. Recall that charged particles bend in a magnetic field, via the Lorentz force. Therefore, the LHC is equipped with 1232 dipole magnets distributed all along the length of the beam pipe to keep the proton bunches turning in the tunnel. The cross section of such a dipole magnet is shown in Fig. 3-2. Each dipole magnet is 14.3m long, weighs 35t, cost nearly 500 KCHF to produce, and has nearly 11 700 amps of current running through it. Only with such massive currents is it possible to generate the appropriate magnetic field strength of 8 T to keep the protons turning. The magnetic field is maintained by titanium-niobium coils, which are kept under cryogenic conditions using liquid helium to achieve the necessary temperature of 1.9 K to reach a superconducting state; this temperature is colder than that of outer space.

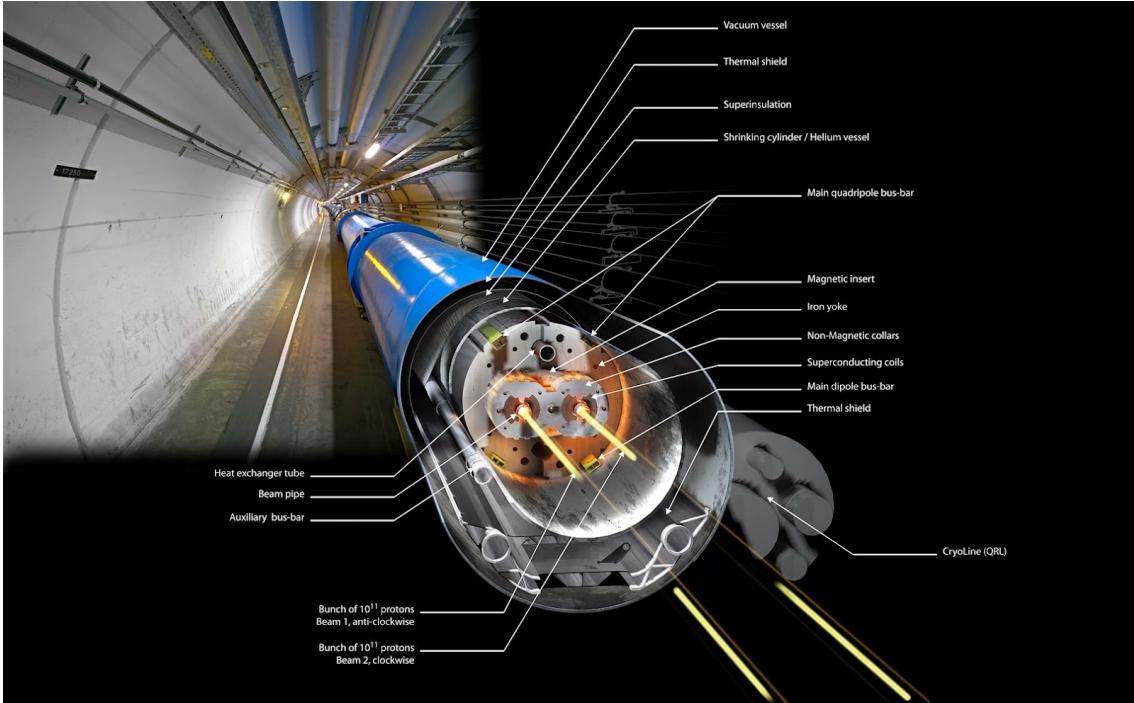


Figure 3-2. A cross section of one of the 1232 dipole magnets which span the entire length of the LHC tunnel.

There are only four specific “Points” along the LHC where the proton bunches actually cross, as shown in Fig. 3-1 (Left, red stars). At each of these four points, there is a unique and gigantic particle detector to catch all the decay products from the pp collisions.

As the two bunches are just about to cross one another, they are squeezed down using quadrupole magnets, focusing the beams more tightly, increasing their chance for tasty pp collisions. During such a bunch crossing (BX), amazingly most of the protons just pass right by one another; out of the possible 100 billion possible collisions that could have occurred, Fig. 3-3 shows that on average only 32 collisions occurred per BX in the LHC 2018 run, according to a particle detector called CMS, described in Chapter 4. It should be mentioned that the luminosity of the LHC is on the order of $L = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$.

As the proton bunches whiz around the LHC, they are given “kicks” from radio-frequency (RF) cavities, which accelerate the protons to a max speed of 99.999 996% c . It is analogous to the timing required when pushing someone on a swing: push at just the right time to increase their momentum. At this speed, *each proton* carries 6.5 TeV of energy, such that a single pp collision

contains a monstrous center-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ —more than enough energy to create new particles like top quarks, Higgs bosons, and potentially BSM particles (Fig. 3-4). In order to analyze such interesting particles, one needs to detect the outgoing particles produced from pp collisions. One such dedicated particle detector is located at Point 5—the Compact Muon Solenoid (CMS) detector—and will be described in the following chapter.

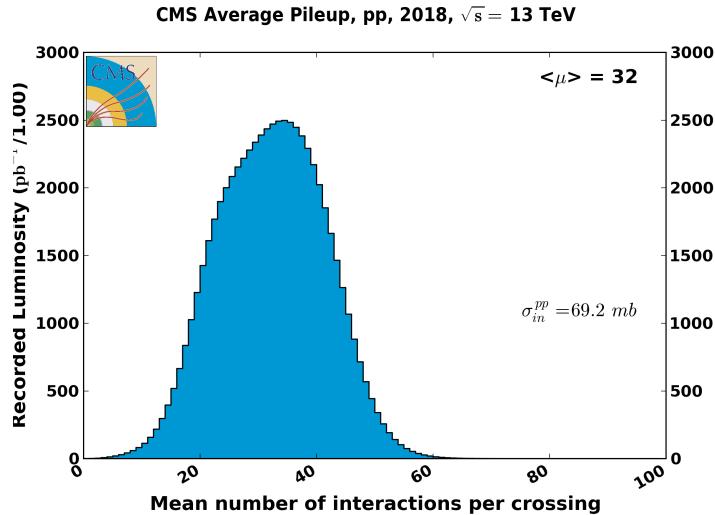


Figure 3-3. Histogram showing the distribution of the average number of pp collisions per proton bunch crossing (pile up) which CMS recorded during the LHC 2018 run.

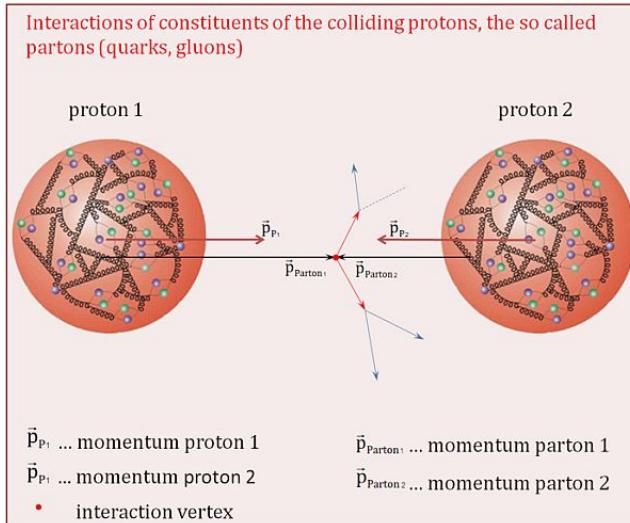


Figure 3-4. Two protons can be smashed together at very high energies to have their constituent partons interact and convert their high energies into new kinds of matter.

CHAPTER 4

THE CMS DETECTOR



Figure 4-1. Life-size poster of the CMS detector, taken during CERN Open Days 2019 in the SX5 warehouse where parts of CMS were assembled.

Weighing in at 14 000 tonnes, standing 15 m (5 stories) tall, and reaching 29 m long, the Compact Muon Solenoid (CMS) experiment is one of the four major particle detectors at the LHC (Fig. 4-1). As shown in Fig. 4-2, CMS is situated approximately 100 m under the earth at the fifth collision point (*Point 5*) along the LHC. In 2012, both CMS and its competing experiment, A Toroidal LHC ApparatuS (ATLAS), independently discovered the Higgs boson.

Recall that the LHC (Chapter 3) directs two proton bunches together on a collision course every 25 ns to produce thousands of new particles per pp collision. These outward-moving particles travel all sorts of directions away from the interaction point. CMS is built around the interaction point in a series of cylindrical subdetectors for nearly hermetic coverage so that most of the particles *must* travel through CMS. The detector sports a solenoid—for which CMS was named—which generates a 3.8 T uniform magnetic field that points longitudinally down the central axis of CMS. This strong magnetic field applies a Lorentz force on the outgoing charged particles, causing them to follow helical, momentum-dependent trajectories. These curved tracks separate from one another which assists in particle identification. Electrically neutral particles experience no Lorentz force and thus travel in straight lines.

The subdetectors measure the properties of the outgoing particles and carefully filter them

out in a clever way (Fig. 4-3). As particles pass through the subdetectors, they interact in a variety of ways. For example, charged particles can leave “hits” in the silicon tracker system and continue through to the next radially outward subdetector, whereas hadronic matter tends to be captured by the hadronic calorimeter. Generally, hits can be reconstructed into particle tracks. From the track curvature, the charge and momentum of the particles can be deduced. Depending on *which* subdetector (or combination of subdetectors) produced signals, the type of particle can be deduced. A few example particles and their associated tracks are shown in Fig. 4-4.

Before discussing each subdetector in the following sections, it is useful to define the coordinate system used in CMS; a typical, right-handed, three-dimensional Cartesian coordinate system (x, y, z) is chosen, whose center $(0, 0, 0)$ is placed at the nominal pp collision point within CMS. The x axis points towards the center of the LHC, the y axis points vertically upward, and the z axis points westward towards the Jura mountains, tangential to the beam direction. Since CMS covers almost the entire spherical 4π steradians around the interaction point, it is convenient to use spherical coordinates (r, ϕ, θ) , in which r measures the radial distance in the x - y plane, ϕ measures the azimuthal angle in the x - y plane as measured from the x axis, and θ measures the polar angle as measured from the z axis. When dealing with ultra-relativistic particles like those produced at the LHC, special-relativistic effects like length contraction must be taken into account and so the coordinate θ becomes frame-dependent. It is thus helpful to convert θ to the Lorentz-invariant quantity called pseudorapidity (η), defined as $\eta = -\ln[\tan(\theta/2)]$.

In the remainder of this chapter, the subdetectors of CMS are described in detail: the silicon tracker in Sec. 4.1, followed by the electromagnetic and hadron calorimeters in Sec. 4.2, then the solenoid and yoke system in Sec. 4.3.

4.1 The Silicon Tracker

100M pixels, 40M pictures per second.

At the heart of CMS is one of the world’s largest silicon detectors: the silicon tracker. The main goal of the silicon tracker is not to capture outgoing particles but to very precisely measure the hits from the charged particles as they pass through it. The tracker also assists in vertex

identification, differentiating between primary and secondary vertices, the latter of which often comes from B meson decays. When multiple pp collisions occur within the same BX (so-called *pile up*), the tracker distinguishes between proton collisions with a resolution of about 100 μm longitudinally and 50 μm transverse to the beam pipe. This is crucial to resolve which outgoing particles came from which pp vertex.

The tracker consists of two types of pure silicon detectors: the pixel detector and the strip detector, each of which is described in turn below.

4.1.1 The Pixel Detector

The innermost part of the silicon tracker is the pixel detector, which is the closest subdetector to the interaction point. The pixel detector is composed of 66 million silicon “pixels”, as shown in Fig. 4-5 (Left, pink). A single pixel is $100\mu\text{m} \times 150\mu\text{m}$ and, collectively, they cover a sensitive area of 1.9 m^2 . Because it sits only 8 cm away from the beam pipe, the pixel detector receives the highest particle flux than any other subdetector: around 10 million particles/ cm^2 per second.

The pixel detector is made of three cylindrical layers and two endcaps that surround the beam pipe. In total, the pixel detector has around 6,000 connections (channels) per cm^2 .

After the LHC Run 1 was completed, the accelerator received luminosity upgrades during the 2013–2014 Long Shutdown. To handle these higher luminosities, the pixel detector was replaced by the CMS Phase-1 pixel detector during the LHC technical stop in 2016–2017. The upgrades outfitted the detector with four barrel layers and three endcap disks per side, which allowed for particle detection up to $|\eta| < 2.5$. The overall mass of the pixel detector decreased and granted the detector with better tracking capability.

4.1.2 The Strip Detector

The outer part of the silicon tracker is called the strip detector, which has 10 million detector strips spread across 10 cylindrical layers. The first 4 layers belong to the tracker inner barrel (TIB) and the remaining 6 layers belong to the tracker outer barrel (TOB), Fig. 4-5 (Left, green and blue, respectively). Both the TIB and TOB have two endcaps associated with them, the TID and TEC, respectively. Accounting for all of its components, the strip detector is sensitive to 200 m^2 .

Fig. 4-6 gives a clearly labelled transverse illustration of the pixel and strip detectors.

4.2 The Calorimeters

4.2.1 Electromagnetic Calorimeter

Overview: Particles that pass through the silicon tracker 4.1 encounter the electromagnetic calorimeter (ECAL). Those particles which interact electromagnetically but not strongly (i.e., mostly photons and electrons) are typically absorbed by the ECAL. The particle's energy is then transferred to the ECAL in the form of an electromagnetic (EM) shower. The size and shape of the EM shower provide information about the particle's energy and trajectory. Since the Higgs boson can decay into two photons, the ECAL played a critical role in detecting this decay mode.

Design: The ECAL is a hermetic, cylindrical, homogeneous sub-detector that consists of a barrel (EB), two endcaps (EE), and a preshower detector in front of each endcap (Figure 4-7, Left). The EB covers $|\eta| < 1.479$ while the EE covers $1.479 < |\eta| < 3.0$. The entire subdetector is composed of transparent lead tungstate (PbWO_4) crystals that point axially towards the center (i.e., towards the interaction point) of CMS. The transparent crystals, one of which is shown in Fig. 4-8 (Left), have a high density (8.28 g/cm^3) which provides the ECAL with radiation resistance and a short radiation length ($X_0 = 0.89 \text{ cm}$). Because so many crystals are used (61 200 crystals in the EB and 7324 in the EE), the ECAL has excellent energy resolution and fine granularity. Each endcap is composed of two *Dees*, one of which is shown in Figure 4-7 (Right). A single Dee carries 3662 crystals. Crystals in the barrel are tapered, having front-face dimensions $2.2 \times 2.2 \text{ cm}^2$, back-face dimensions $2.6 \times 2.6 \text{ cm}^2$, and are 23.0 cm long ($25.8 X_0$). Crystals in the endcaps are also tapered, with front-face dimensions $2.862 \times 2.862 \text{ cm}^2$, back-face dimensions $3.0 \times 3.0 \text{ cm}^2$, and are 22.0 cm long ($24.7 X_0$). This gives a single crystal from the barrel a volume of approximately 132.5 cm^3 (i.e., 132.5 mL)—about the volume of a small cup of coffee—yet it has a surprisingly hefty mass of 1.5 Kg [8].

Physics: When electrons or photons pass through the ECAL, they create an EM shower. Electrons radiate more photons as they accelerate around PbWO_4 nuclei, in a process called *bremssstrahlung*. Meanwhile, near the presence of a nucleus, high-energy photons pair produce

into e^+e^- . This cycle of electron/photon production disperses all the initial particle energy into a spray of decreasingly lower-energy particles until it all runs out; this is the EM shower.

The ECAL crystals then scintillate (emits photons) in proportion to the amount of energy deposited by the interacting particle. The scintillator photons are detected by avalanche photodiodes on the back of each barrel crystal or by vacuum phototriodes in the endcap crystals (Fig. 4-8, Right). After 1 BX (25 ns), approximately 80% of the scintillated light is emitted.

An energy deposit in the ECAL could come from either an electron or a photon. In order to tell the difference, information from the silicon tracker is used. Charged particles, like electrons, will leave hits in the tracker and follow a curved path, whereas photons are electrically neutral and thus will not show any signs within the silicon tracker. So long as the tracker and ECAL communicate effectively with each other, then they help distinguish between electrons and photons. Charged hadrons interact only minimally with the ECAL, instead continuing on to the Hadron Calorimeter. Neutral hadrons can be detected by the ECAL preshower near the ECAL endcaps which helps distinguish a single photon from π^0 mesons as they decay into two photons with a narrow opening angle, making it look as if the two photons are a single photon. The preshower detector allows CMS to distinguish between collimated, low-energy diphoton pairs and single high-energy photons.

From the original spray of particles that leave the interaction vertex, the short-lived particles have decayed into lighter, more stable, particles and the ECAL has filtered out most of the electrons and photons. The remainder of the spray is comprised of both hadronic matter and muons, however the hadrons are many times more numerous than the muons and so are filtered out next.

4.2.2 Hadron Calorimeter

Overview: The particles that survive the ECAL—typically only muons and hadrons—then enter the hadron calorimeter (HCAL). Its primary purpose is to absorb the hadronic matter emerging from the interaction point and to measure the corresponding jet energies. The absorbed jets cause the HCAL to scintillate photons which are then converted into electrical signals. These

signals help deduce the original jet energies and missing transverse energy (E_T^{miss}) from the event.

Design: Dissimilar to the ECAL (Sec. 4.2.1) in material composition but similar to it in shape, the HCAL is a brass cylindrical scintillator. Although it has a barrel (HB) and two endcaps (HE), it has two more detectors than the ECAL: the outer calorimeter (HO) and the forward calorimeter (HF). The HB spans the pseudorapidity range $|\eta| < 1.3$, the HE spans $1.3 < |\eta| < 3$, and the HF spans $3 < |\eta| < 5.2$, as shown in Fig. 4-9. With a thickness of over 1 m, the HB is sandwiched between the barrels of the ECAL and the solenoid (Sec. 4.3) at radial values $r = 1.77$ m and $r = 2.95$ m, respectively. Because the HB and HE are located within the solenoid's strong magnetic field of 3.8 T, they were both constructed out of a non-magnetic absorber called *C26000 cartridge brass*. This absorber has a density of 8.53 g/cm³ and an interaction length (λ_I) of 16.42 cm. The thickness of the HB increases as $1/\sin\theta$ so that at $|\eta| = 0(1.3)$ the absorber thickness is 5.82(10.6) λ_I . The HB is composed of two half-barrels, where each half-barrel is built from 18 identical azimuthal wedges and each wedge spans 20°. Each wedge is divided into four ϕ segments so that a single ϕ segment spans $\Delta\phi = 0.087$.

Since the volume available to the HCAL is so limited—and in order to stop any particles that might traverse the entire HCAL and solenoid—the HO (the *tail catcher*) is situated outside the barrel of the solenoid. The HF is located 11.2 m from the interaction point. All tiles within a single ϕ segment are grouped together into a single tray unit. The scintillator is also segmented into 16 η sectors, the first(last) of which is located at $|\eta| = 0(1.3)$. This way each tile covers $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$. Each layer has 108 trays.

Physics: Since hadrons are the only particles to interact via the strong force, the HCAL is designed to have a high nuclear density. This ensures ample opportunity for hadrons to radiate gluons and convert with the Similal to the ECAL, the HCAL will scintillate in proportion to the amount of energy of the captured particle. The incoming hadrons will *hadronize* (i.e., produce a hadronic shower), generating jets of quarks and gluons which are bound in various ways forming protons, neutrons, pions, kaons, etc. Interestingly, the HCAL is made using over a million old, brass shell casings from the Russian Navy back from World War II.

About 34% of the particles produced from LHC pp collisions enter the HE region, so the HE was built to handle high rates (MHz).

The entire HCAL utilizes approximately 70,000 plastic scintillator tiles. The active material in the HB is 3.7-mm-thick Kuraray SCSN81 plastic scintillator, selected for its radiation hardness and long-term stability. Hadron showers \rightarrow tiles scintillate \rightarrow scintillated photons are collected by 0.94-mm-diameter green double-cladded wavelength shifting (WLS) fibers (Kuraray Y-11), which carry the light to hybrid photodiodes (HPD).

4.3 The Solenoid and the Steel Return Yoke

The Compact Muon *Solenoid* sports one of the world's most energetic solenoids which is paramount to the success of CMS. Particles that exit the HCAL (subsec. 4.2.2) arrive at the cylindrical magnet which is 12.5 m in length, has a bore diameter of 6 m (6.3 m when cold), and generates a uniform 3.8 T magnetic field parallel to the beam line. To produce such a large and uniform magnetic field inside the approximately 360 m^3 volume (Fig. 4-10), an 18 000 amp current travels through the 4-layer, superconducting, NbTi coils. This magnetic field stores a massive 2.6 GJ of energy—approximately the kinetic energy of an Airbus A320 in flight—and is 100 000 times stronger than Earth's magnetic field, as measured on the surface. The magnet has such a large stored-energy-to-cold-mass ratio (11.6 KJ/Kg) that it experiences a physical deformation of 0.15% while being energized.

As charged particles travel through any magnetic field, they experience a magnetic (Lorentz) force perpendicular to their direction of travel. The Lorentz force (\vec{F}_B) exerted on a particle with charge q depends on the particle's velocity (\vec{v}) and the magnetic field (\vec{B}) via the cross product

$$\vec{F}_B = q\vec{v} \times \vec{B}.$$

Since the force is necessarily perpendicular to the velocity, the resulting trajectory is helical. Projecting the helix on the x - y plane (since the magnetic field points in the $+z$ direction) allows the particle tracks to typically be separated from one another. Each track has a corresponding

radius of curvature (R) which relates to its transverse momentum (p_T) through

$$p_T = qBR.$$

The relative change in p_T (i.e., the *momentum resolution*) is given by

$$\frac{\delta p_T}{p_T} \propto \frac{p_T}{BL^2}. \quad (4-1)$$

Steel Return Yoke: Most of the mass of CMS comes from the *steel return yoke* which helps to redirect the magnetic field back on itself. The yoke system constitutes 10 000 tonnes, which is 89% of the total mass of CMS. It is comprised of 5 wheels and 2 endcaps.

4.4 The Muon System

Although it is the farthest system from the interaction point, the muon system is one of the most important within the Compact *Muon* Solenoid detector. Of the particles emerging from the interaction point, electrons and photons are absorbed by the ECAL (subsec. 4.2.1) and hadronic matter is absorbed by the HCAL (subsec. 4.2.2). This filtration process leaves only neutrinos and muons to enter the muon system, which is the outermost detector situated past the solenoid (sec. 4.3). As mentioned in , neutrinos are the only weakly-interacting, electrically-neutral SM particles which makes them incredibly difficult to detect directly. In fact, neutrinos interact with normal matter so little that a light-year (9.46 trillion Km) of lead would only stop half of the neutrinos moving through it. Thus, the detection of neutrinos produced from pp collisions is inferred via E_T^{miss} on a per-event basis. Muons, on the other hand, have a mass of 105.7 MeV (relatively heavy for a weakly-interacting, electrically-charged particle) and live a billion-billion times longer than a Higgs boson: the average lifetime of a muon is $\tau_\mu = 2.2 \times 10^{-6}$ s. These properties are what determined the properties of the muon system within CMS to consist of its four main subdetectors, each of which is described in the following subsections:

1. CSC (cathode strip chambers, subsec. 4.4.1),
2. DT (drift tubes, subsec. 4.4.2),

3. RPC (resistive plate chambers, subsec. 4.4.3),

4. GEM (gas electron multiplier, subsec. 4.4.4).

4.4.1 Cathode Strip Chambers

Overview: A cathode strip chamber (CSC) is a multi-wire proportional chamber capable of precisely measuring the position of muons which enter and ionize the gas within. Spatial coordinates are obtained by the collection of electrical signals along the cathode strips (ϕ), anode wires (r), and across multiple CSC layers (z). CSCs are found exclusively on the two endcaps of CMS, with each endcap bearing 270 chambers (Fig 4-11). The chambers are arranged azimuthally around the beam pipe in four disks per endcap allowing for contiguous coverage in ϕ . In total, the CSC system provides an effective detection area of about 5,000 m², has a total gas volume that exceeds > 50 m³, and contains over 400,000 read-out channels.

Design: Each CSC is trapezoidal in shape with its narrow end pointed toward the beam pipe (Fig. 4-14, Left). The chambers are arranged in rings and each chamber subtends either 10° or 20° in ϕ , as described in subsubsec. 4.4.1.1. The CSCs cover the pseudorapidity range of $0.9 < |\eta| < 2.4$ (Fig. 4-12). A single CSC is composed of six layers (or *gas gaps*), each of which is filled with a carefully prepared gaseous mixture¹ of Ar:CO₂:CF₄ (Fig. 4-13). The gas mixture flows from one layer to the next in a zigzag manner and has a flow rate of ????. Within every layer, the gas mixture surrounds approximately 80 copper strips, each of which spans radially away from the interaction point. A single strip is about 8.4 mm wide at the narrower end of the CSC, about 16 mm at the wider end, and is separated from its neighboring strip by about 0.5 mm. Per layer, the inner gas also surrounds over 1,000 gold-plated tungsten wires, which are oriented azimuthally (so approximately orthogonal to the strips). Each wire is approximately 50 μm in diameter and separated from its neighboring wire by about 3.2 mm. A collection of 16 consecutive wires forms a *wire group*, which is about 5 cm wide and creates a single anode read-out channel. A single wire plane has five independently controlled HV segments (Fig. 4-14, Right). The largest CSCs

¹The gas mixture ratio of Ar:CO₂:CF₄ was chosen to be 5:4:1 to maximize the lifetime of the CSCs as they endure radiation damage through years of use. The CO₂ is used as a non-flammable quencher to reach even larger electron multiplicities, while the CF₄ helps prevent polymerization along (aging of) the wires.

are 3.4 m long as measured along a strip and 1.5 m wide as measured along a wire.

Physics: As a muon passes through a CSC layer, it has the opportunity to interact with and ionize an Ar atom in the gas mixture. The wires are under high voltage (3,600 V) which causes the ionized electron to accelerate towards the positively-charged strips. The accelerating electron collides with and ionizes Ar atoms along its path toward the strip. This liberates even more electrons, thus forming an *electron avalanche* (Fig. 4-15). The total number of ionized electrons per initial electron is referred to as the *multiplicity* (or *gas gain*), which can reach as high as 100,000.

The electron avalanche is collected by a cathode strip and becomes an electrical signal. This signal is processed by the cathode front-end boards (CFEBs). The Ar^+ ions similarly distribute a charge signal onto the negative wires. The cluster of charge that arrives at a strip is more widely spread than the charge which arrives along a wire. Therefore, comparator logic is implemented on the strips to narrow down the precision to the order of 100 μm by using half-strip information (Fig. 4-16).

The muon passes through the next 5 layers of the CSC, further ionizing the gas mixture and generating electrical signals along the wires and strips. A signal on a wire provides an r -coordinate, on a strip provides a ϕ -coordinate, and through the other layers provides a z -coordinate. Taken together, this information helps to reconstruct the three-dimensional trajectory of the muon.

The number of hits recorded by the CSC will determine if an event was significant enough to be worth saving. If so, then its precise positions on the wires and strips will be read out by the Data Acquisition (DAQ) system and be stored for further data analysis. When a CSC is taking live data it can resolve approximately 2 mm in r - ϕ , whereas during offline analysis the resolution improves by a factor of more than 20: the ME1/1 and ME1/2 chambers can resolve distances as small as 75 μm in r - ϕ , while the other chambers can resolve 150 μm . It is worth noting that a CSC can accommodate up to 1 KHz/cm².

Gas gap is 9.5 mm.

4.4.1.1 CSC Numbering Scheme

The two endcaps are labelled as “ME+” and “ME-”, depending on whether they are situated in the $+z$ direction or $-z$. Both endcaps are structurally symmetric, so it is sufficient to discuss only one in detail. The ME+ endcap has four disks: ME+1 is the first disk and the one closest to the interaction point, while ME+4 is the fourth and the farthest away. Within each disk, there are either two or three “rings” of CSCs, as shown in Fig. 4-12 (green). These rings are labelled as $\text{ME}+D/R$, where D indicates the disk number and R indicates the ring number. For example, ME+2/1 refers to the second disk and the first ring (the ring closest to the beam pipe). All rings contain 36 CSCs, except for $\text{ME}\pm X/1$, for $X = 2, 3, 4$, which contain only 18 CSCs. Finally, the CSCs are given one final number to label them on the ring: the CSC that sits along the positive x -axis in the coordinate system of CMS is given the number “01”, e.g., ME+4/2/01. The CSCs are then numbered incrementally following the positive azimuthal direction.

4.4.2 Drift-Tube Chambers

Overview: Functionally similar to but structurally different from a CSC (subsec. 4.4.1), a DT is a collection of gaseous detector cells (Fig. 4-17). A single DT cell has an anode wire and two cathode strips and operates on the same principle as a CSC, providing timing and position measurements of muons (Fig. 4-18). While CSCs are found only on the endcaps of CMS, drift-tube chambers (DTs) are placed exclusively along the barrel (Fig. 4-19). The DT system is therefore composed of concentric cylindrical stations, with the central axis parallel to the beam pipe. Altogether, there are 250 DTs built inside of, between, and outside of the iron yoke.

Design: The first three stations contain 60 DTs each, while the station farthest from the beam pipe contains 70 DTs. The DTs are distributed among 5 wheels, 4 stations, and 12 sectors within the muon barrel (MB) system which uses the following numbering scheme: $\text{MB}/W/A/S$, where W is the wheel number (-2 to 2), A is the station number (1 to 4), and S is the sector number (1 to 12). This accounts for only $5 \times 4 \times 12 = 240$ chambers, so the remaining 10 DTs are found in station 4, sectors 4 and 10, in each wheel (Fig. 4-19).

Each tube cross section is $13 \times 42 \text{ mm}^2$, within which a single anode wire is made of

gold-plated stainless-steel that has a diameter of 50 μm . Different voltages are applied to the anode wires (+3600 V), electrode strips (+1800 V), and cathode strips (-1200 V). The gas mixture within a DT is similar to that of a CSC, containing approximately 85% Ar + 15% CO₂.

Physics: DTs are a good choice for the barrel because of the low rate and low strength of magnetic field. Uniform drift field of 1.5 KV/cm. A high p_{T} muon track will cross all 4 DT stations within the pseudorapidity range of $|\eta| < 0.8$. The reconstruction efficiency of such a track is better than 95%.

A single SL has a time resolution of only a few nanoseconds which provides excellent bunch crossing identification. To assist in determining muon p_{T} and timing, the DT system delivers the following muon track segment information to the L1 trigger: the position of the center of gravity (to a precision of 1.5 mm) and the corresponding angle (to a precision of 20 mrad), w.r.t. the SL reference frame. The total resolution of a DT in $r\text{-}\phi$ is about 100 μm which is comparable to the deviation caused by multiple scattering, for a muon with $p_{\text{T}} \leq 200 \text{ GeV}$.

The gas gain is comparable to that of a CSC (about 100,000).

The maximum drift distance in a DT is 21 mm, as can be seen in Fig. 4-18, which corresponds to a drift time of 380 ns. This value

4.4.3 Resistive Plate Chambers

Overview: Even though CSCs 4.4.1 and DTs 4.4.2 cover the entire pseudorapidity range of $0 < |\eta| < 2.4$, redundancy is important. Therefore, resistive plate chambers (RPCs) are placed in both the endcaps and barrel to provide excellent timing and spatial resolution, comparable to that of scintillators, to supplement the measurements of the CSC and DT systems. An RPC is a gaseous parallel-plate detector

RPC has better time resolution than the 25 ns between two contiguous pp bunch crossings.

Design:

Gas mixture comprised of 96.2% C₂H₂F₄ (1,1,1,2-tetrafluoroethane) + 3.5% iC₄H₁₀ (isobutane) +

The RPC system consists of 480 rectangular chambers that are oriented parallel to the beam

pipe. A single RPC is 2.455 mm long

Physics:

4.4.4 Gas Electron Multipliers

Overview:

Design:

Physics:

(GEM)

Located in the forward region. Filled with Ar/CO₂ gas mixture. The first GEMs At the time of this dissertation writing,

4.5 Trigger System

4.5.1 The Level-1 Trigger

Trigger system helps to manage the enormous collision rate at Size of 1 event is about 3 MB. This would be TB of data per second, which isn't feasible. Two stages of a trigger system: Level 1 Trigger (L1) and the High Level Trigger (HLT).

L1 Trigger has about a trigger rate of 100 KHz as a hard limit. It's latency is only 3.2 microsec.

At the start of a fill in the LHC, the trigger rate is maximum. INSERT L1 trigger rate plot.

In order to remain at the maximal L1 trigger, prescales are implemented which switch on new triggers, while turning off others. Collision data which passes certain triggers can be “parked” on disk and analyzed later (FIXME e.g., during Long Shutdowns) Data parking is used to Data scouting: The HLT has a trigger rate of about 1.5 KHz. It can write out about 5 GB/s to disk and has a much higher latency of 0.5 s compared to the L1 Trigger.

4.5.2 The High-Level Trigger

Only record 1K events/s some events only occur only once per 10B collisions!

Use the GIF of the m4l plot with growing Data.

4.6 Object Reconstruction

A single pp collision yields thousands of particles, each of which must be identified by the CMS detector in order to be used in physics analyses. The primary software for object identification within CMS is called Particle Flow. Electron and photon objects are reconstructed (Sec. 4.6.1) using the silicon tracker (Sec. 4.1) and ECAL (Sec. 4.2.1) systems. The reconstruction of muon objects (Sec. 4.6.2) uses information from all subdetectors, the most important of which are the silicon tracker and muon system (Sec. 4.4). Tau objects are the final type of charged leptons to be reconstructed (Sec. 4.6.3), which may decay hadronically—thus leaving hadronic energy signatures in the HCAL—and by the silicon tracker detecting a displaced secondary vertex, thanks to the relatively long lifetime of the tau lepton. Hadronic matter is grouped into jet objects, each of which is carefully reconstructed (Sec. 4.6.4.) using information from the silicon tracker, ECAL, and HCAL (Sec. 4.2.2). Finally, after accounting for all the momenta and energies of the observed particles, the missing transverse energy (E_T^{miss}) is evaluated (Sec. 4.6.5).

4.6.1 Electron and Photon Reconstruction

Electrons and photons both leave energy deposits within the crystals of the ECAL—so how are the two kinds of particles differentiated from one another? Since photons are electrically neutral, the silicon tracker will not detect them. Another consequence of their neutrality is that photons do not bend in a magnetic field; their energy deposits will point radially back to the vertex from which they originated. On the other hand, electrons *are* charged and so will not only be detected within the layers of the silicon tracker but will also reveal a curved track. Hence, the energy deposits from electrons within the ECAL will not necessarily point towards the vertex from which the electron came [12].

Within CMS, track fitting and pattern recognition typically occur within a single framework called the Combinatorial Track Finder (CTF), which is an extension of the Kalman filter [13]. However, the reconstruction of tracks of charged particles is computationally difficult the CTF As electrons pass through matter, they may radiate away some of their energy in the form of bremsstrahlung photons. The distribution of this energy loss is extremely non-Gaussian and is not

accurately modeled by the Kalman Filter. Instead, the CMS tracker uses a Gaussian-sum filter (GSF) algorithm [14] to model the energy-loss distribution as a mixture of Gaussian functions. This ultimately improves the momentum resolution of electrons.

4.6.2 Muon Reconstruction

The most reliable muon detection occurs when both the silicon tracker *and* the muon system both register hits corresponding to the same particle object. When both subsystems detect such a particle—most likely a muon—it is called a *global muon*. This can be compared to the situation when only the muon system registers hits, in which case this is most likely a *standalone muon*. The opposite case occurs when only the tracker registers hits and the corresponding particle is termed a *tracker muon* [15].

4.6.3 Tau Reconstruction

4.6.4 Jet Reconstruction

4.6.4.1 Jet Energy Correction

4.6.4.2 Tagging b-Jets

4.6.5 MET Reconstruction

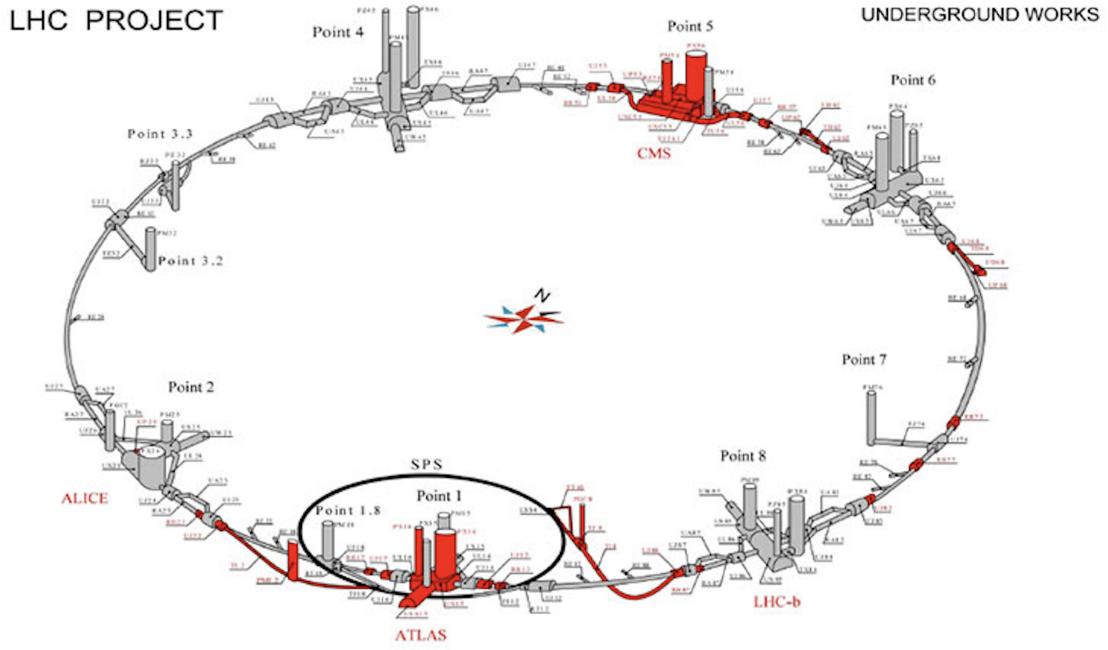


Figure 4-2. Points of interest along the LHC (Points 1–8). Collisions occur at Points 1 (ATLAS), 2 (ALICE), 5 (CMS), and 8 (LHCb), whereas the remaining points are used for LHC beam maintenance and testing.

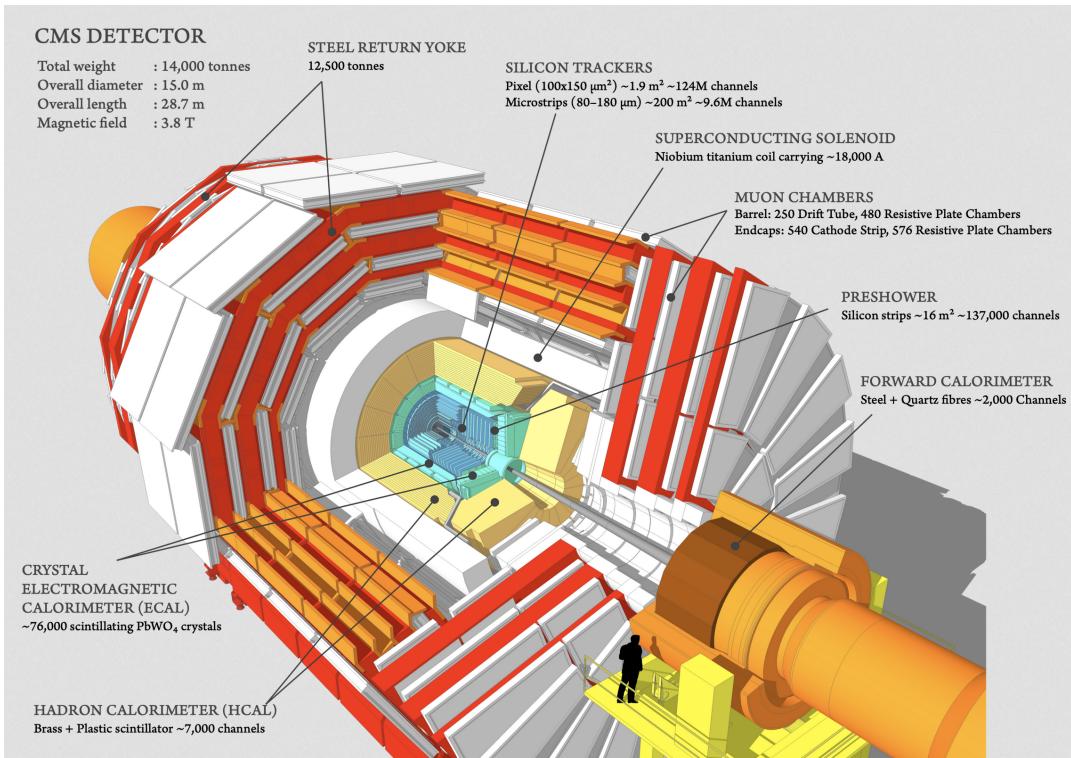


Figure 4-3. Cut out of the CMS detector showing its various subdetector components.

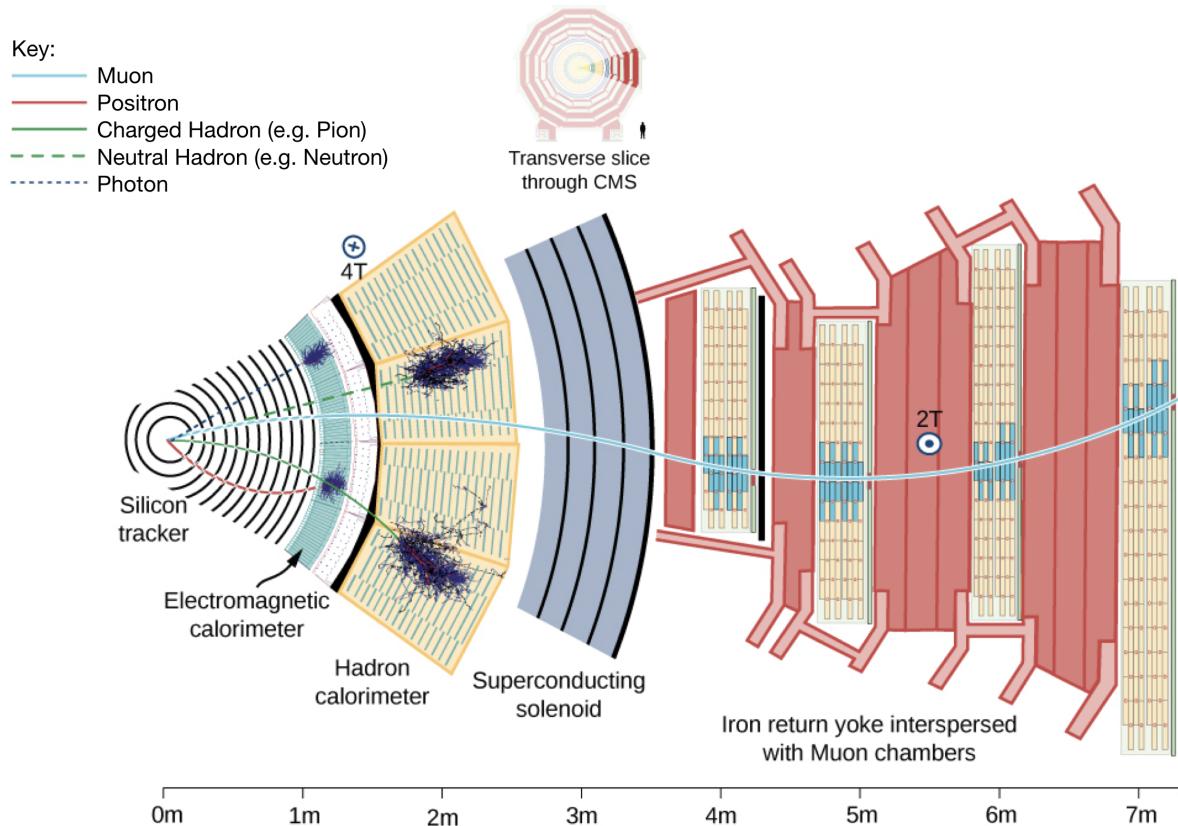


Figure 4-4. A transverse view of CMS showing the “filtration process” as different particles pass through different subdetectors. A positron (solid red line) curves due to the presence of the magnetic field and gets stopped in the ECAL, creating an EM shower. A photon (blue dashed line) does not get detected at all by the Silicon Tracker, since it has no electric charge. It continues through to the ECAL and makes a shower here, like the positron. Charged hadrons (solid green line) will show curved tracks from the Silicon Tracker, may leave some trace in the ECAL, but primarily get stopped by the HCAL creating hadronic showers. Neutral hadrons (dashed green line) do not interact with the tracker, and only undergo EM showers a little in the ECAL, but show most energy deposits in the HCAL. Muons (solid blue line) are detected by the Silicon Tracker and then mostly pass through the other subdetectors without interacting until they finally reach the Muon System. Using the Lorentz force law and knowing which direction the magnetic field is pointing, one can deduce the sign of the charge of the particle. Based on the radius of curvature from the trajectory, one can then calculate the momentum and energy of the particle.

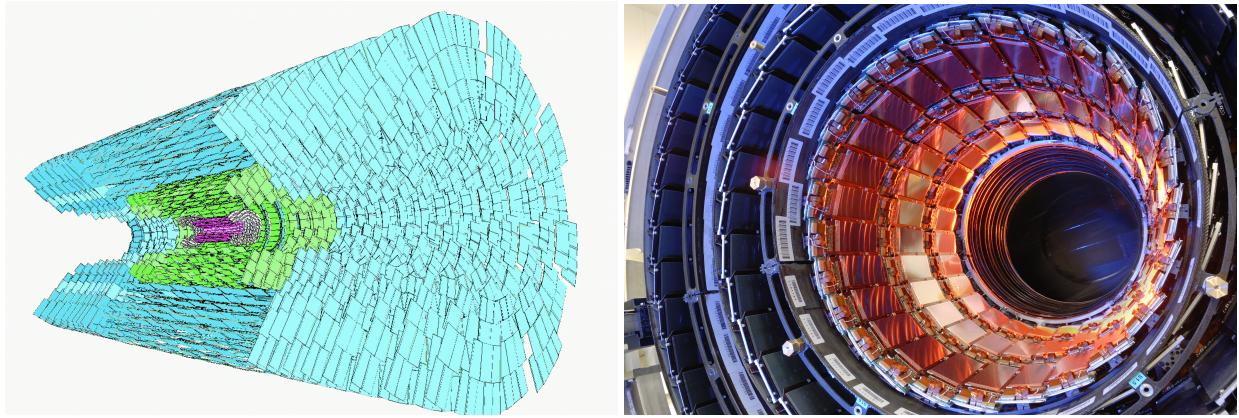


Figure 4-5. (Left) A simulation of the silicon tracker, showing the 3 cylindrical layers of the pixel detector (pink), 4 layers of the TIB (green), and the 6 layers of the TOB (blue) of the strip detector. The endcap components are also shown. (Right) A picture of the real silicon tracker at the center of CMS.

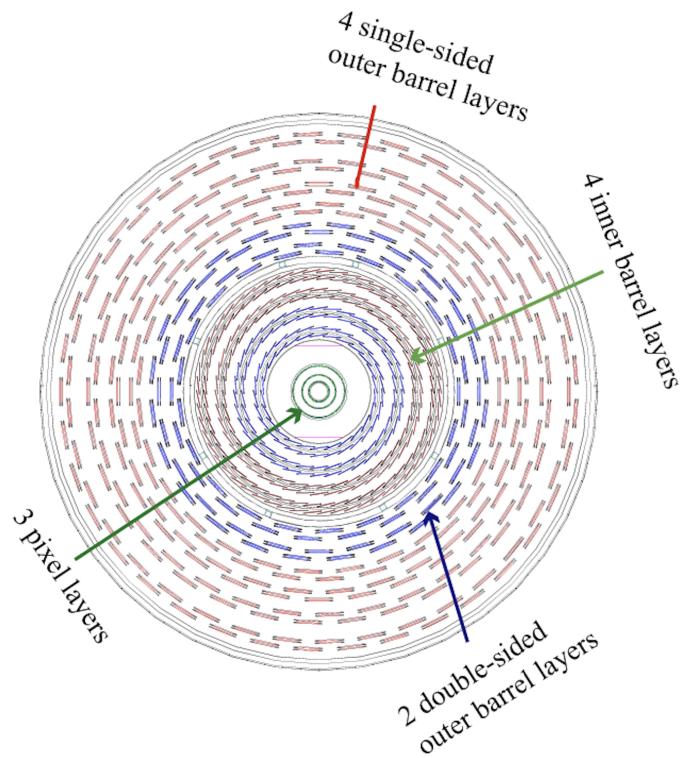


Figure 4-6. A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved.

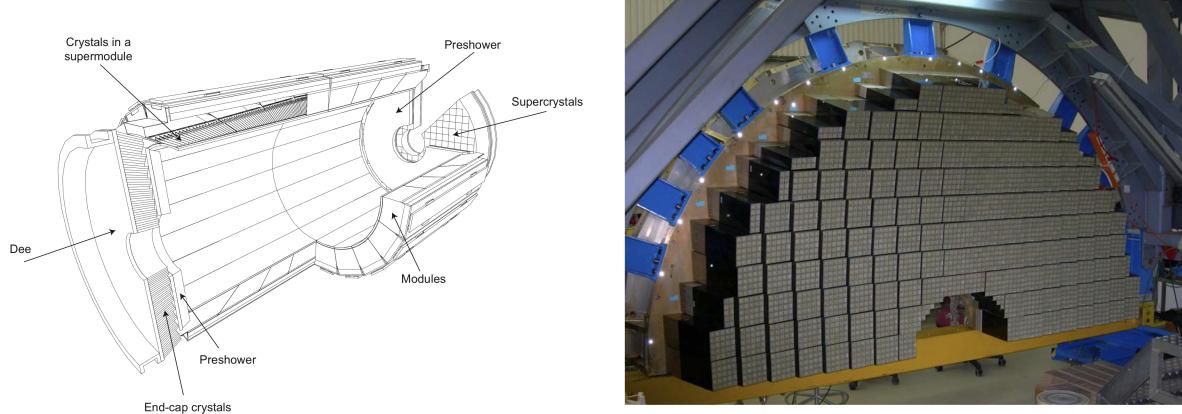


Figure 4-7. (Left) Cross-sectional view of the electromagnetic calorimeter of CMS. (Right) One of the Dees which comprise the EE. Each square of 5×5 crystals constitutes a “super-crystal”. Figure taken from Ref. [11].



Figure 4-8. (Left) ECAL crystals made from PbWO₄ are grown in a lab. (Right) Although made mostly of metal, ECAL crystals are transparent and have a photomultiplier detector attached at the end.

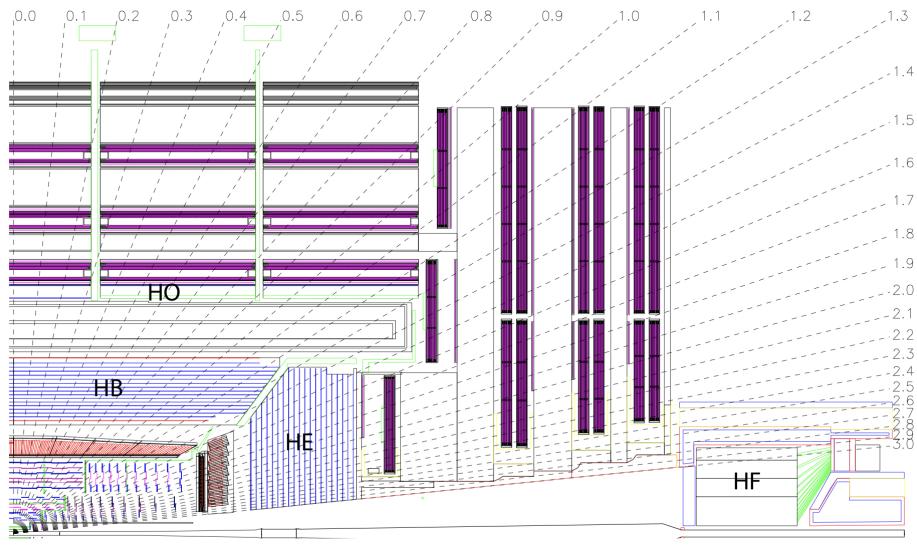


Figure 4-9. A cross-sectional quadrant view of CMS showing the locations of the HCAL components: the barrel (HB), outer (HO), endcap (HE), and forward (HF) detectors.

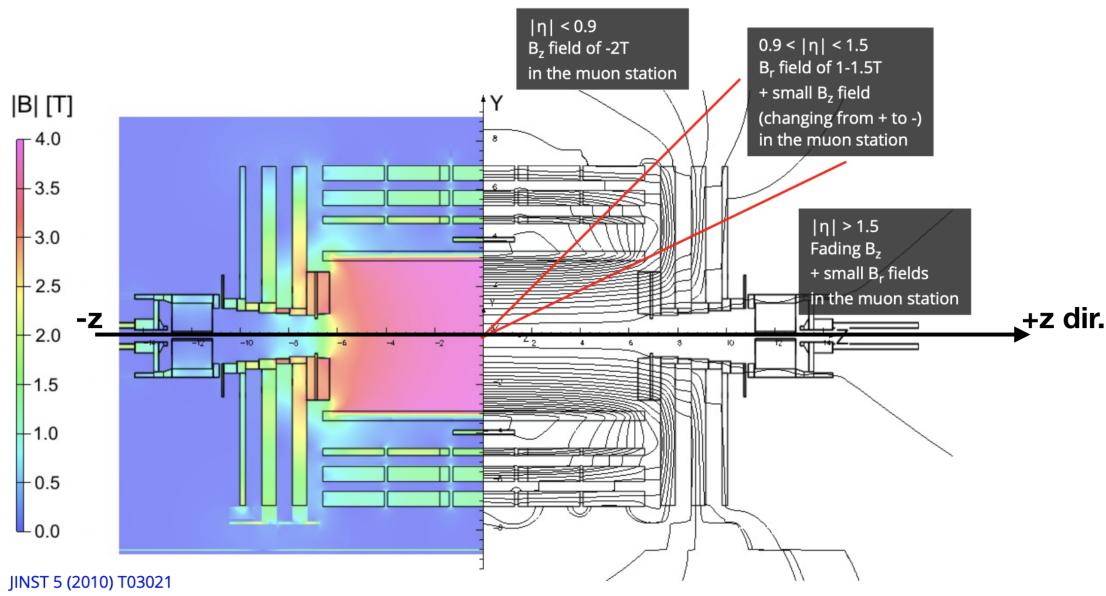


Figure 4-10. A longitudinal cross section of CMS showing the values of the magnetic field over the volume of CMS and various field lines. The magnetic field reaches its maximum of 3.8 T in the center of the detector.

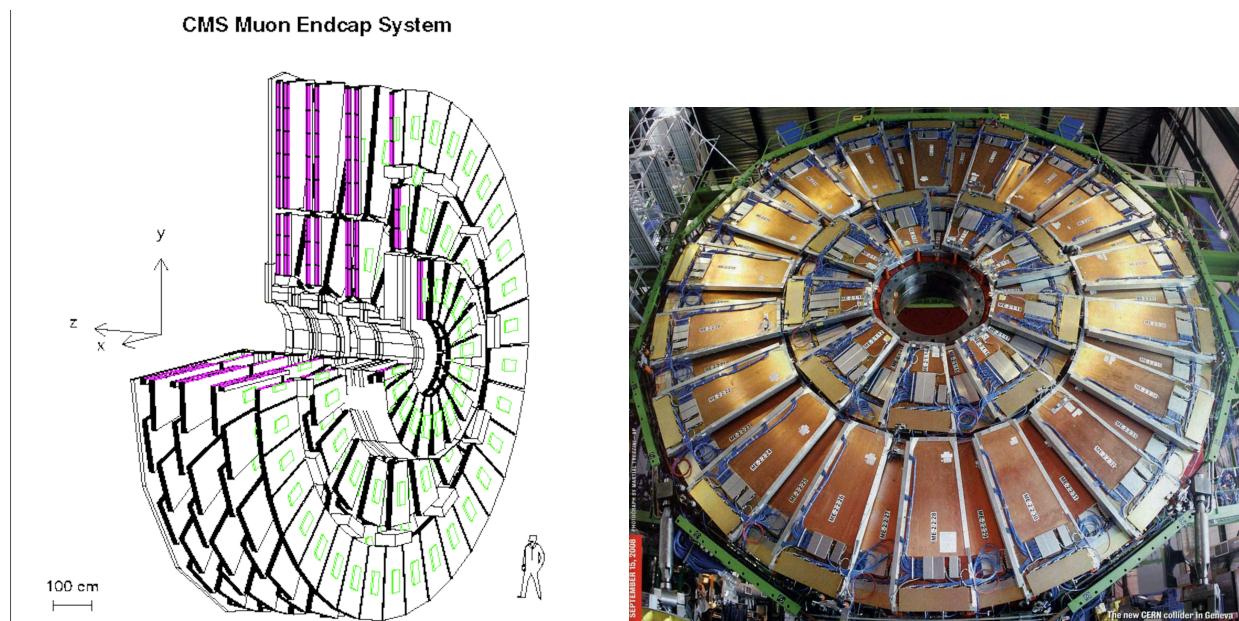


Figure 4-11. (Left) A simulated cut out view of the ME+ endcap, with the coordinate system of CMS in the top-left corner. (Right) The actual ME-2 disk of CMS is shown, revealing its ME-2/1 and ME-2/2 rings of CSCs.

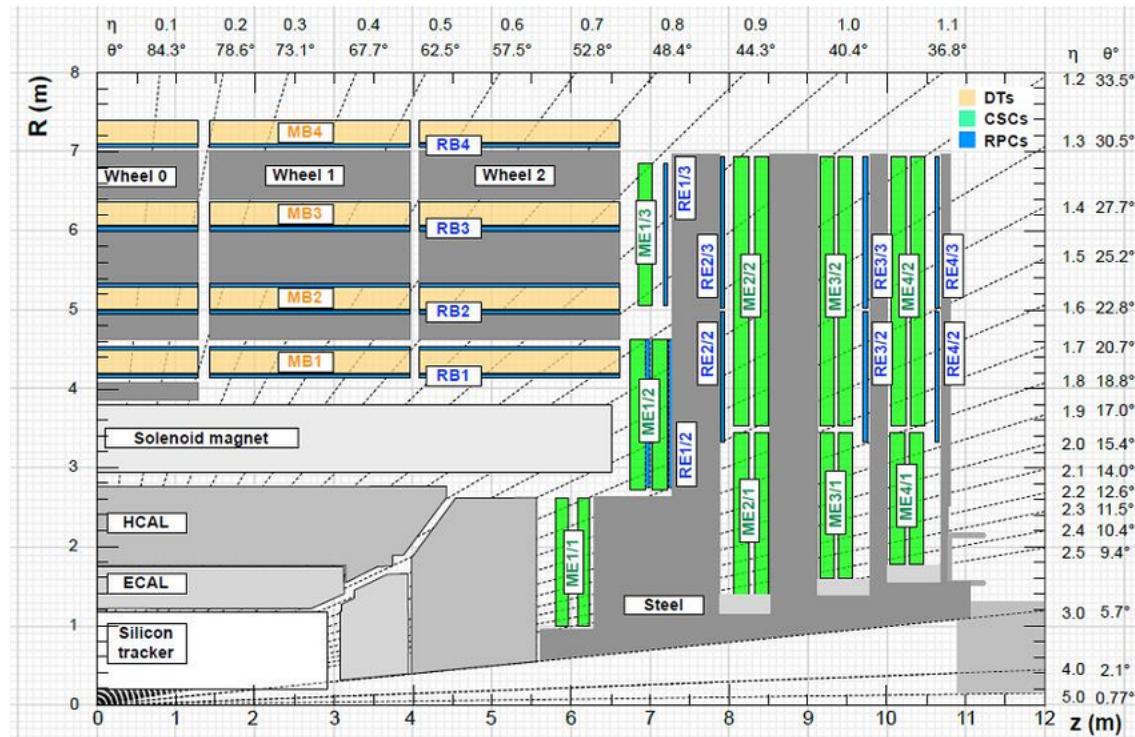


Figure 4-12. Longitudinal cross section of CMS, showing the different pseudorapidity values (η) and also the different subdetector regions.

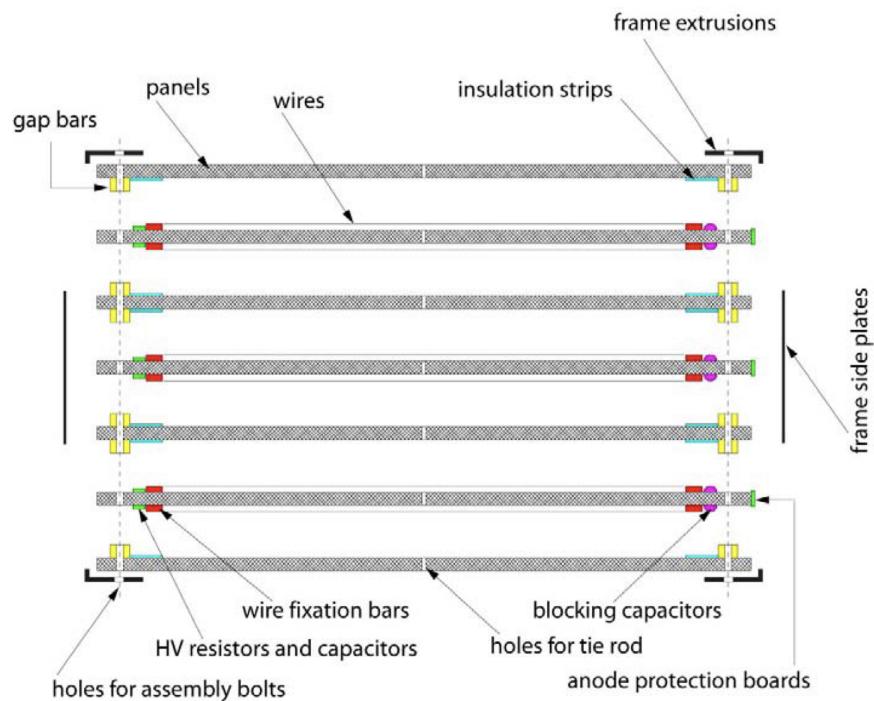


Figure 4-13. Exploded view of the cross section of a CSC showing how the 7 panels come together to form the 6 gas gaps between the panels.

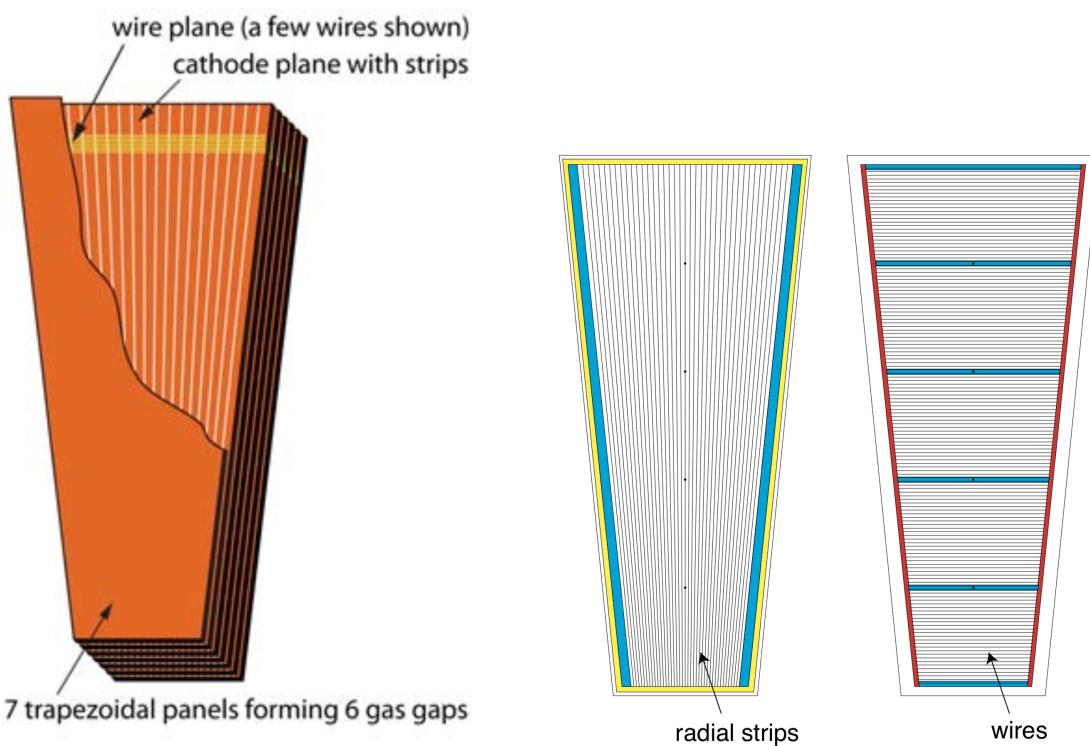


Figure 4-14. (Left) A CSC with its top layer exposed. You can see very thin gold-plated tungsten wires which actually span the entire width of the CSC. Thicker vertical strips run along the length. (Right) More detail showing the radial strips and the horizontal wires. Also shown is the 5 segments of a CSC.

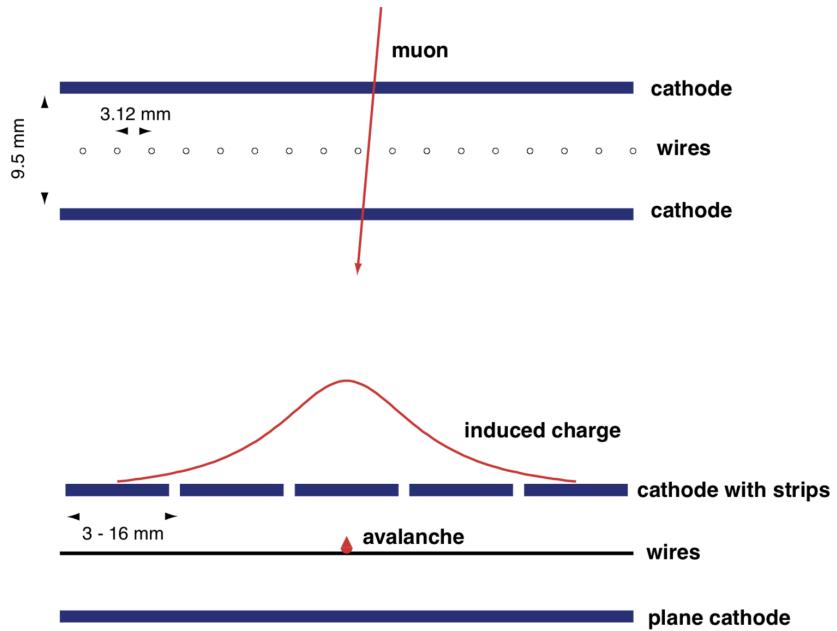


Figure 4-15. A muon passes through one of the gaseous layers of the CSC, ionizing the gas mixture and inducing a charge on the anode wires and cathode strips.

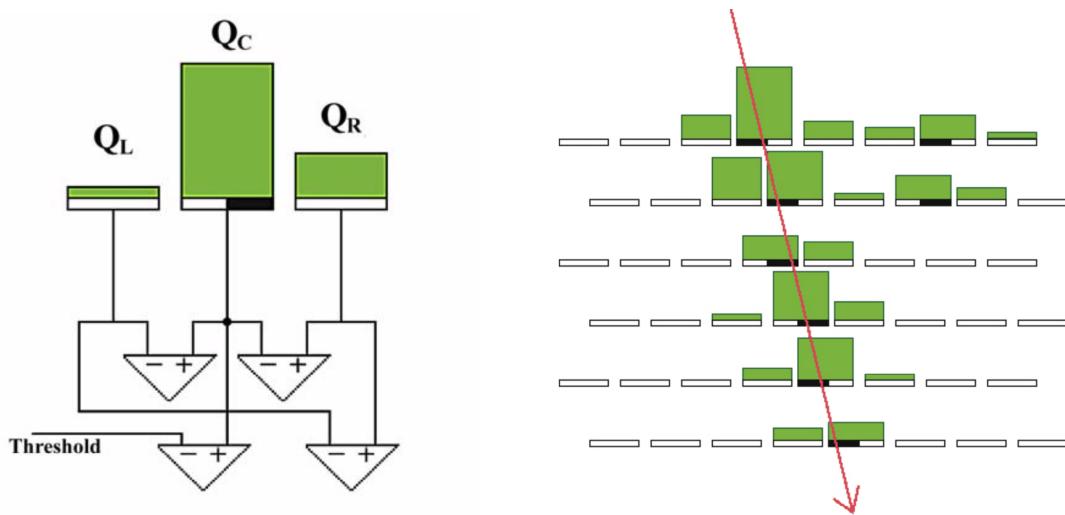


Figure 4-16. (Left) Comparators are used to compare neighboring strip cluster charge to determine on which half-strip the peak charge resided. (Right) A muon passes through all six layers of a CSC inducing charge on various half-strips.

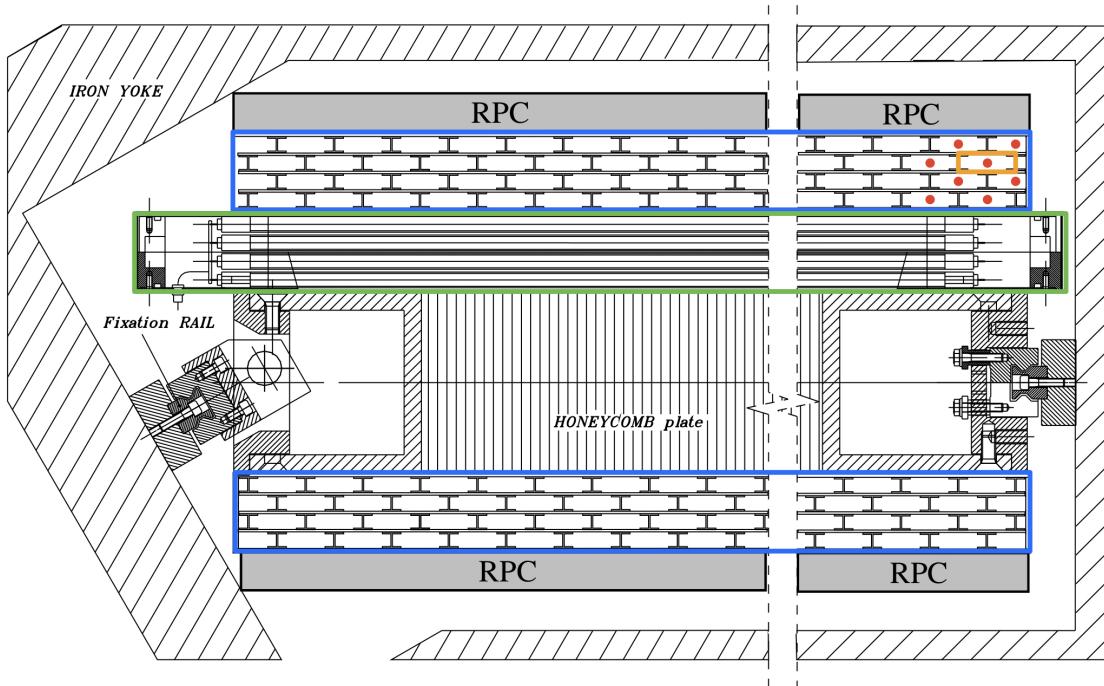


Figure 4-17. A drift-tube chamber with 3 superlayers (SLs) is sandwiching a honeycomb plate. Two SLs are indicated by the blue rectangles, whose cells (orange rectangle) have anode wires (red dots) oriented parallel to the beam pipe. The third SL is indicated by the green rectangle and is staggered orthogonally to the other two SLs.

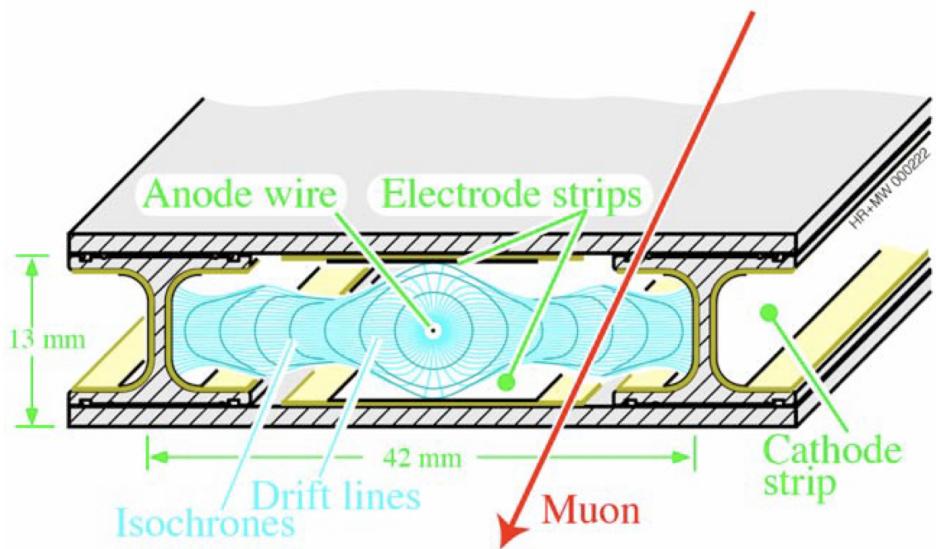


Figure 4-18. A cross section of a DT cell showing the drift lines (light blue), isochrone lines (dark blue), dimensions of the cell, and locations of the anode wire and cathode strips.

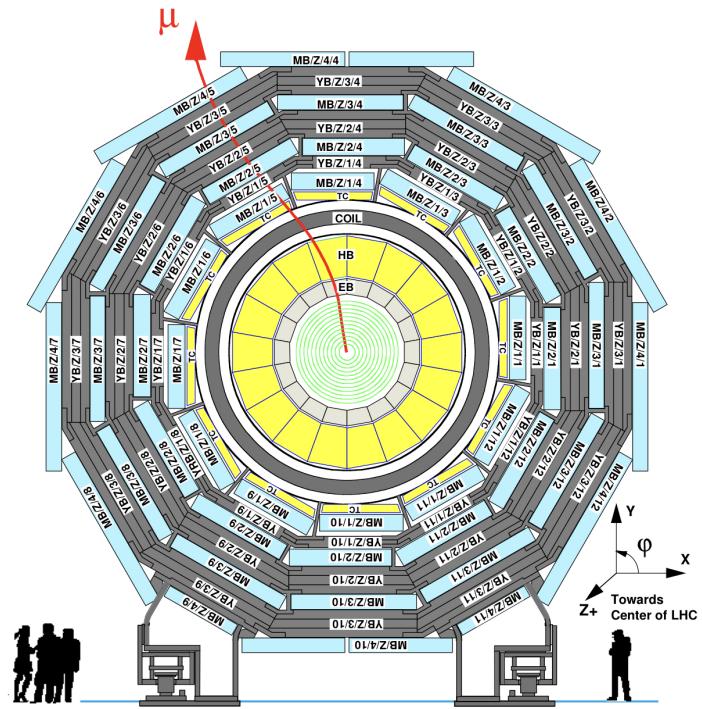


Figure 4-19. A cross section of CMS showing the locations and numbering scheme of the drift tubes in the barrel.

CHAPTER 5

HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL

5.1 Motivation

When the CMS and ATLAS collaborations announced the discovery of the Higgs boson on July 4, 2012, this was a momentous achievement in particle physics because the so-called “missing” piece of the SM was found. Evidence of the Higgs boson’s existence also motivates the associated Higgs field, which permeates all of spacetime and explains the origins of the masses of all the other massive fundamental particles (Chapter 2).

The Higgs boson is interesting for a variety of reasons. First, it is currently one of a kind—the only fundamental scalar particle ever discovered at the time of this writing. Second, the mass of the Higgs boson theoretically determines the stability of our very Universe (Fig. 1-2). Third, the unique boson could be a portal to new physics—i.e., physics beyond the Standard Model (BSM)—e.g., by decaying into BSM low-mass dilepton mass resonances (Chapter 6). Fourth, the Higgs boson may not be the only one of its kind; some BSM models theorize that other kinds of Higgs bosons may exist. Fifth, *are we certain that the Higgs boson discovered in 2012 is the same as the one predicted by the SM?* To check this, it is necessary to compare the Higgs boson’s measured properties to its predicted ones. One such property is the mass of the Higgs boson (m_H).

This chapter details the measurement of m_H full Run 2 data from the LHC as analyzed by the CMS detector. Although many previous measurements of m_H have already been made (e.g., by the ATLAS and CMS collaborations as shown in Fig. 5-1), the measurement presented in this dissertation gives the world’s most precise value of m_H .

First, a general overview of the logic and analysis workflow for the m_H measurement is motivated in Sec. 5.2. The specific data sets, simulated samples, and triggers used in the analysis are then detailed in Sec. 5.3. Then the event reconstruction and event selection are described in Sec. 5.4. Afterwards, an analysis of the background estimation is given in Sec. 5.5. The signal modeling and improvements utilized in this measurement are then laid out, which include the kinematic discriminant, per-event mass uncertainties, and the vertex constraint in Sec. 5.7. A treatment of the systematic uncertainties follows in Sec. 5.8. The chapter concludes with a summary of the m_H measurement results in Sec. 5.9.

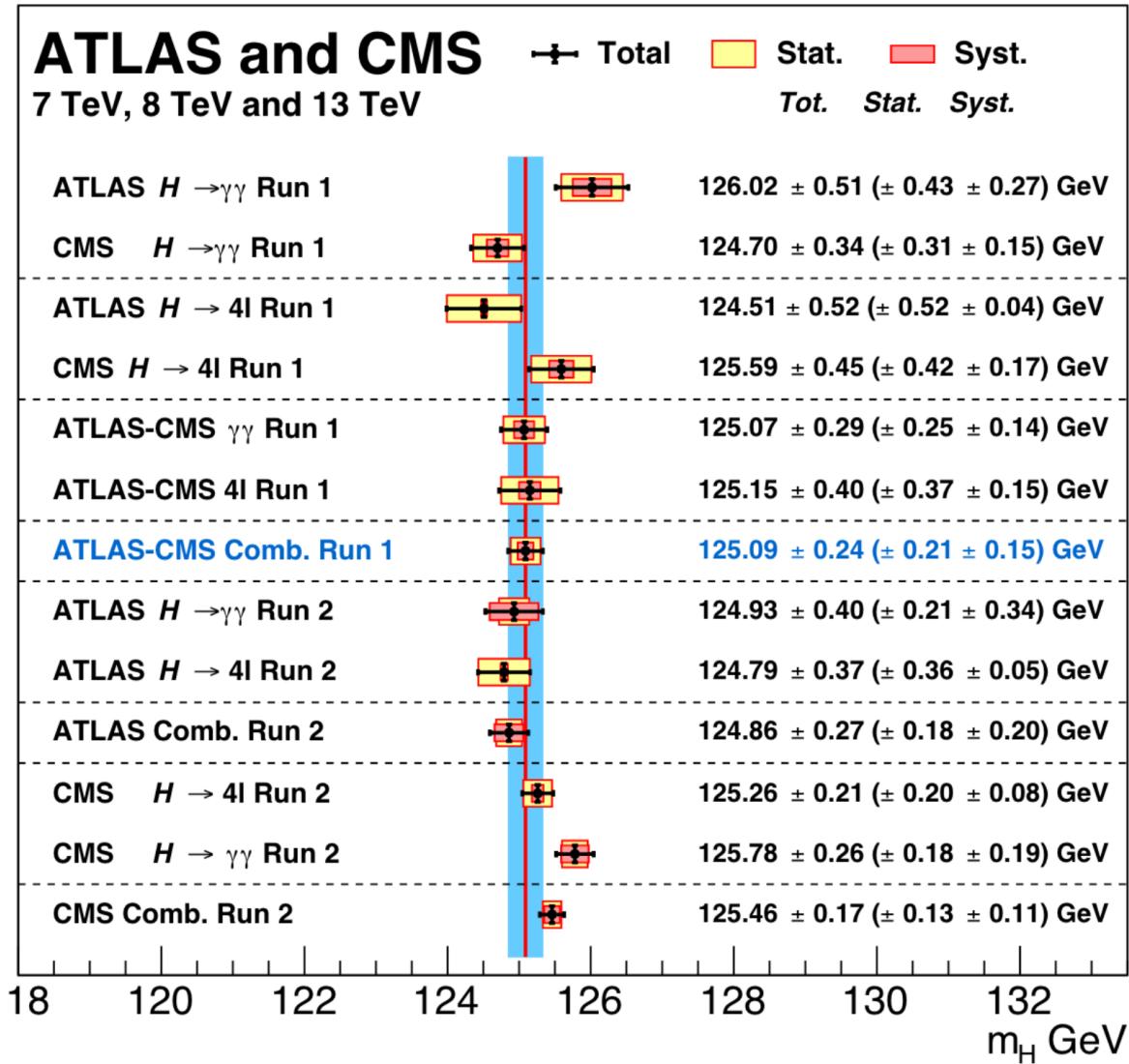


Figure 5-1. Various measurements of m_H made by the CMS and ATLAS collaborations in the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels, in Runs 1 and 2 of the LHC. Plot taken from [8].

5.2 Analysis Overview

The first step to performing a precision measurement of m_H is to “observe” many Higgs bosons. However, production of a Higgs boson is essentially nonexistent in everyday conditions and is still extremely rare even in the high-energy pp collisions of the LHC (Chapter 3). At a center-of-mass energy of 13 TeV, the total inclusive inelastic cross section of two protons colliding is 70 mb. Comparing this cross section to the cross section of Higgs boson production ($\sigma_{pp \rightarrow H} = 59 \text{ pb}$) shows that a Higgs boson is produced in approximately one out of every *billion* pp collisions—a rare event indeed.

To complicate matters further, the Higgs boson has a *very* short mean lifetime of only $1.6 \times 10^{-22} \text{ s}$. Thus, the existence of the Higgs boson is not *directly* detected by CMS (Chapter 4) but is instead *inferred* from its stable decay products which enter the various subdetectors. Among all the fundamental particles so far discovered, the Higgs boson has the second largest mass (approximately 125 GeV), which gives the scalar boson sufficient energy to decay into at least 9 different final states. Each decay mode occurs with a different probability—the *branching fraction* or *branching ratio* (\mathcal{B})—whose value depends on m_H , as shown in Fig. 5-2. The question then becomes, “*Which Higgs boson decay mode is the most useful for the measurement of m_H ?*”

Owing to its large signal-to-background ratio of approximately 2 and its relatively rare four-lepton final state, the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel is chosen and is called the *signal process*. On average, a Higgs boson will decay into two Z bosons (one on-shell and the other off-shell) only 2.6% of the time. In turn, each Z boson *may* decay into two opposite-sign, same flavor (OSSF) leptons ($Z \rightarrow \ell^+ \ell^-$, where $\ell = e, \mu$) on average approximately 6.7% of the time. This signal process then gives rise to four distinct final states: 4e, 4μ, 2e2μ, 2μ2e. The branching ratio for the overall signal process is then calculated as:

$$\mathcal{B}(H \rightarrow ZZ^* \rightarrow 4\ell) = \mathcal{B}(H \rightarrow ZZ^*) [\mathcal{B}(Z \rightarrow \ell^+ \ell^-)]^2 = 1.8 \times 10^{-3}.$$

Thus, a signal event is expected to be produced only once in about every *trillion* pp collisions.

The strategy is then to search the pp collision data collected and analyzed by CMS for all the

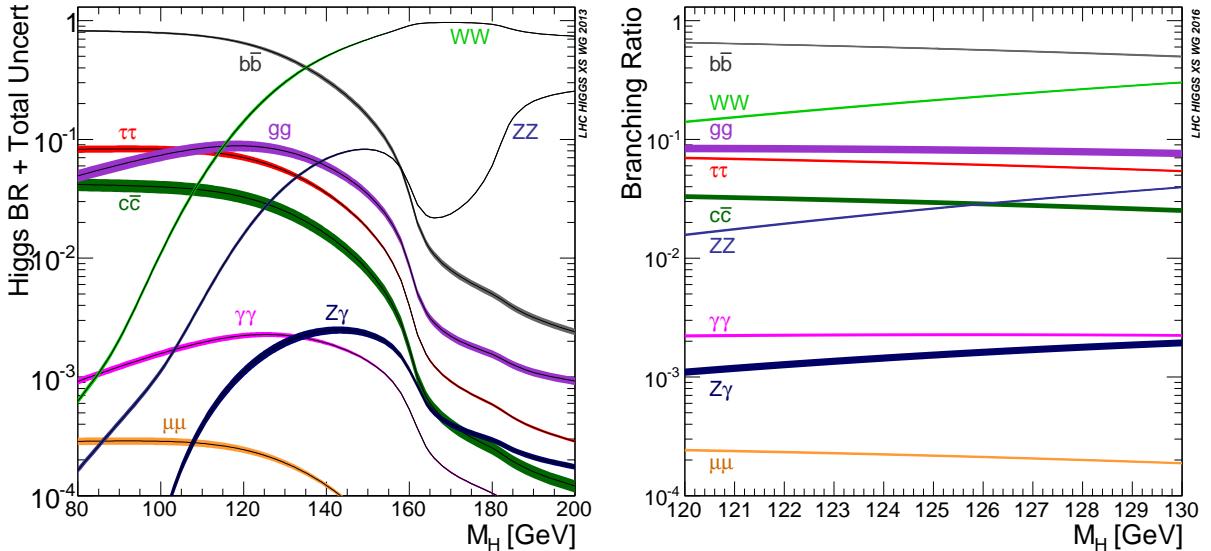


Figure 5-2. The branching ratios of various Higgs boson decays as a function of the Higgs boson mass over a wide range (Left) and a narrow range (Right) of values.

detected $H \rightarrow ZZ^* \rightarrow 4\ell$ events. The task is not so straightforward; events in the data are categorized—not by the entire decay process—but by their final state, based on which triggers fired to collect which events. For each chosen event, the subdetectors of CMS (Sects. 4.1–4.4) provide a plethora of track and energy-detection information to reconstruct *objects*, which are representations of the underlying particles within the event. The reconstructed objects are then assembled in a fashion that checks if the decay logic coincides with the signal process of interest: $H \rightarrow ZZ^* \rightarrow 4\ell$. For example, a pair of OSSF lepton-like objects should appear to come from a Z-like object—i.e., having a nominal mass of approximately 91 GeV and zero net electric charge—instead of, say, appearing to come from a H-like object. Two such Z-like objects must be formed and should appear to come from a H-like object. All throughout, the reconstructed event must obey all physics conservation laws (energy, momentum, charge, etc.) Furthermore, the associated objects are often required to pass certain detector selection criteria (e.g., $p_T^\mu > 5$ GeV). These criteria are analysis-specific and are collectively called the *event selection* of the analysis.

Although the event selection is devised to ideally select only signal events, it is not guaranteed; there are certain physics process that have exactly the same initial and final states as

the signal process. Such processes “contaminate” the collected signal events and are called *background processes*. Concretely, Fig. 5-3 shows how identical initial-state gluons can react to produce exactly the same final-state particles, while producing different intermediate particles: the signal process (Left), initiated by gluon-gluon fusion vs. a background process (Right) which skips the intermediate Higgs boson. It is imperative for particle physics analyses to maximize the number of collected signal events while minimizing the number of collected background events. Section 5.5 discusses the associated background processes and how to estimate the number of events these contribute to the signal region.

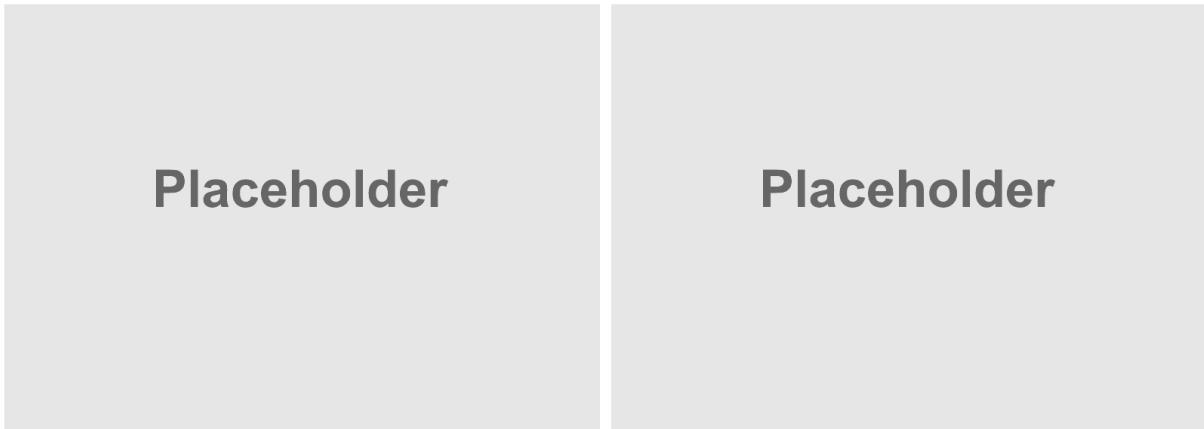


Figure 5-3. Feynman diagrams showing how the initial and final states are the same for the signal process ($gg \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$, Left) and one possible background process ($gg \rightarrow ZZ^* \rightarrow 4\ell$, Right).

Before drawing conclusions from the data themselves, it is necessary for particle physicists to make predictions about their analysis using simulated events or *simulation*. Such events simulate a specific process (e.g., $pp \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$), governed by some theoretical framework that is programmed mathematically into a software package. Computer programs like `MADGRAPH5_amc@NLO` and `POWHEG` can simulate millions of rare (or even *fictitious*) events in just a single day, whereas gathering the same number of events in CMS data might otherwise take *decades*! Furthermore, software can even simulate the particles as they travel through the simulated detectors. Programs like `GEANT4` show analysts what to expect as the particles interact with a virtual version of the CMS detector. Predictions from simulation can then be compared to

the truth—*the data*—as a way to check the accuracy of the analysis. For example, a surplus of events in data where none was expected may lead to the discovery of new particles, as was the case for the discovery of the Higgs boson.

So how is the measurement of m_H obtained? Since the signal process is $H \rightarrow ZZ^* \rightarrow 4\ell$, conservation of energy leads one to expect that $m_{4\ell} \approx m_H$. Although this is not how the final measurement is obtained, it is a reasonable starting point. The distribution of $m_{4\ell}$ values reveals that the Higgs boson mass resonance stands well above the distribution of expected background events, as shown in Fig. 5-4. Simulated signal events are then used to predict the *line shape* of this signal peak. This signal modeling is performed using a double-sided Crystal Ball function to fit the line shape, for various mass points of m_H , in each of the four final states.

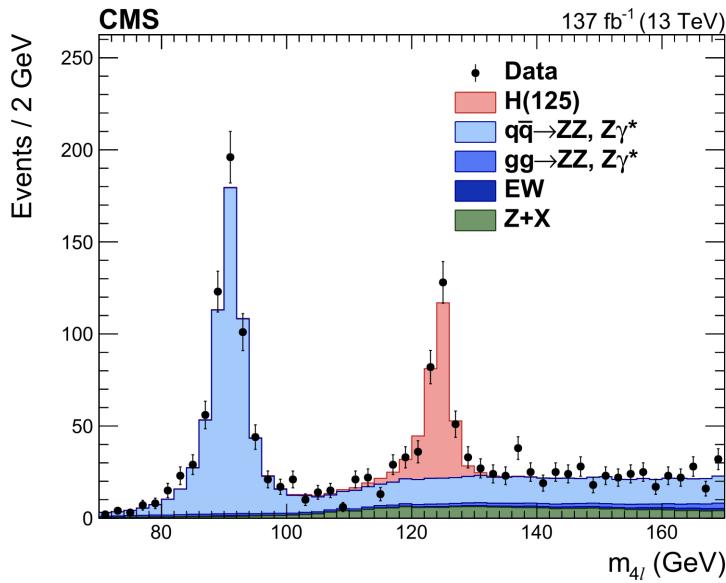


Figure 5-4. Distribution of $m_{4\ell}$ from $H \rightarrow ZZ^* \rightarrow 4\ell$ events using Full Run 2 data. Plot taken from .

The precision on the measured value of m_H is improved by implementing several techniques: calculating a per-event matrix element kinematic discriminant, deriving correction factors for the uncertainty on $m_{4\ell}$ within various regions of phase space, reevaluating the lepton p_T values using a Z_1 mass constraint, and constraining the four muon tracks to the selected vertex (*vertex constraint*). The implementation of these techniques is discussed in Sec. 5.6.

Important in any scientific analysis is the careful study of all the associated uncertainties that inherently come with making *any* measurement. There are two main kinds of uncertainties: the first is *statistical* uncertainty, which depends on how many data points were collected to make the measurement, while the second is *systematic* uncertainty, which depends on the precision of the instruments used to make the measurement. The systematic uncertainties for this analysis include:

- L_{int} (2.5%).
- Lepton identification and reconstruction efficiencies (2.5–9%).
- Lepton energy scale.
- Estimation of the reducible background (40%).

The systematic theoretical uncertainties include:

- Renormalization and factorization scale uncertainties.
- Choice of the set of parton distribution functions.
- Branching fraction uncertainties for signal and background processes.

All of these ingredients are used in the final measurement, which is obtained via a likelihood fit on the m_H spectrum to extract the most likely value of m_H . The results of the precision measurement of m_H are summarized in Sec. 5.9.

5.3 Analyzed Data

The CMS experiment recorded LHC 3 data during 2016, 2017, and 2018 (collectively called Run 2), corresponding to an integrated luminosity of 137.1 fb^{-1} . These events were categorized by the trigger system 4.5 into different data sets, depending on which triggers “fired”, i.e., whether the event passed the trigger criteria or not. The names of the data sets are listed in Tables 5-4–5-6 and follow the format `/object_type/campaign/datatier`. This analysis uses the ultra legacy (UL) reconstruction []. It should be noted that the 2016 data are split into 2 different reconstruction versions, starting at Run2016F: the first version is “pre-VFP”, which uses HIP mitigation (HIPM) in the reconstruction, and the second is “post-VFP”, which uses the default reconstruction.

For the $H \rightarrow ZZ^* \rightarrow 4\ell$ analysis, Sec. 5.3.1 lists the triggers used, Sec. 5.3.2 details the data sets used, and Sec. 5.3.3 summarizes the simulated samples used.

5.3.1 Triggers

Table 5-1. Trigger paths used to collect 2016 data (pre- and post-VFP) for the measurement of m_H .

HLT path	Prescale	Primary data set
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v*	1	DoubleEG
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*	1	DoubleMuon
HLT_TripleMu_12_10_5_v*	1	DoubleMuon
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_v*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Ele25_eta2p1_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele27_eta2p1_WPLoose_Gsf_v*	1	SingleElectron
HLT_Ele27_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele32_eta2p1_WPTight_Gsf_v*	1	SingleElectron
HLT_IsoMu20_v* OR HLT_IsoTkMu20_v*	1	SingleMuon
HLT_IsoMu22_v* OR HLT_IsoTkMu22_v*	1	SingleMuon
HLT_IsoMu24_v* OR HLT_IsoTkMu24_v*	1	SingleMuon

5.3.2 Data Sets

5.3.3 Simulated Events

The data set names containing the simulated events for signal and background processes are listed in Tables 5-7–5-9.

5.3.4 PileUp reweight

Simulated samples are reweighted taking into account pileUp distribution. PileUp weight is extracted for each year matching simulation and data distribution. The minimum bias cross-section used for each year is 69.2 mb. PileUp distributions for each year are shown in Fig. 5-5.

Table 5-2. Trigger paths used to collect 2017 data for the measurement of m_H .

HLT path	Prescale	Primary data set
HLT_DoubleEle33_CaloIdL_MW_v*	1	DoubleEG
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v*	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*	1	DoubleMuon
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v*	1	DoubleMuon
HLT_TripleMu_10_5_5_DZ_v*	1	DoubleMuon
HLT_TripleMu_12_10_5_v*	1	DoubleMuon
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ_v*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_DZ_v*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Ele35_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele38_WPTight_Gsf_v*	1	SingleElectron
HLT_Ele40_WPTight_Gsf_v*	1	SingleElectron
HLT_IsoMu27_v*	1	SingleMuon

Table 5-3. Trigger paths used to collect 2018 data for the measurement of m_H .

HLT path	Prescale	Primary data set
HLT_DoubleEle25_CaloIdL_MW_v*	1	DoubleEG
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	DoubleEG
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*	1	DoubleMuon
HLT_TripleMu_10_5_5_DZ_v*	1	DoubleMuon
HLT_TripleMu_12_10_5_v*	1	DoubleMuon
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_DZ_v*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_DZ_v*	1	MuonEG
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v*	1	MuonEG
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*	1	MuonEG
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*	1	MuonEG
HLT_Ele32_WPTight_Gsf_v*	1	SingleElectron
HLT_IsoMu24_v*	1	SingleMuon

Table 5-4. Names of the UL 2016 data sets and the corresponding integrated luminosities (L_{int}).

Data set name	Integrated luminosity (fb $^{-1}$)
/DoubleEG/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v3/MINIAOD	
/DoubleMuon/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016C-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016C-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016C-HIPM_UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016C-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016C-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016D-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016D-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016D-HIPM_UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016D-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016D-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016E-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016E-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016E-HIPM_UL2016_MiniAODv2-v5/MINIAOD	TODO
/SingleMuon/Run2016E-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016E-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016F-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016F-HIPM_UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016F-HIPM_UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016F-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016F-HIPM_UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016F-UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016F-UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016F-UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016F-UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016F-UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016G-UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016G-UL2016_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2016G-UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016G-UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016G-UL2016_MiniAODv2-v2/MINIAOD	
/DoubleEG/Run2016H-UL2016_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2016H-UL2016_MiniAODv2-v2/MINIAOD	
/SingleElectron/Run2016H-UL2016_MiniAODv2-v2/MINIAOD	TODO
/SingleMuon/Run2016H-UL2016_MiniAODv2-v2/MINIAOD	
/MuonEG/Run2016H-UL2016_MiniAODv2-v2/MINIAOD	

Table 5-5. Names of the UL 2017 data sets and the corresponding integrated luminosities (L_{int}).

Data set name	Integrated luminosity (fb^{-1})
/SingleMuon/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	TODO
/DoubleEG/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	TODO
/DoubleEG/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	TODO
/DoubleEG/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2017E-UL2017_MiniAODv2-v2/MINIAOD	TODO
/DoubleEG/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	
/SingleElectron/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	
/DoubleMuon/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	TODO
/DoubleEG/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	

Table 5-6. Names of the UL 2018 data sets and the corresponding integrated luminosities (L_{int}).

Data set name	Integrated luminosity (fb^{-1})
/SingleMuon/Run2018A-UL2018_MiniAODv2-v3/MINIAOD	
/DoubleMuon/Run2018A-UL2018_MiniAODv2-v1/MINIAOD	TODO
/EGamma/Run2018A-UL2018_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2018A-UL2018_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2018B-UL2018_MiniAODv2-v2/MINIAOD	
/DoubleMuon/Run2018B-UL2018_MiniAODv2-v1/MINIAOD	TODO
/EGamma/Run2018B-UL2018_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2018B-UL2018_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2018C-UL2018_MiniAODv2-v2/MINIAOD	
/DoubleMuon/Run2018C-UL2018_MiniAODv2-v1/MINIAOD	TODO
/EGamma/Run2018C-UL2018_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2018C-UL2018_MiniAODv2-v1/MINIAOD	
/SingleMuon/Run2018D-UL2018_MiniAODv2-v3/MINIAOD	
/DoubleMuon/Run2018D-UL2018_MiniAODv2-v1/MINIAOD	TODO
/EGamma/Run2018D-UL2018_MiniAODv2-v1/MINIAOD	
/MuonEG/Run2018D-UL2018_MiniAODv2-v1/MINIAOD	

Table 5-7. Names of simulated signal and background samples for 2016 data.

[1] “RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM”
 or
 “RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM”

Name of signal data set	$\sigma \times \mathcal{B}$ (pb)
GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV709_pythia8/[1]	0.01333521
VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV709_pythia8/[1]	0.001038159
WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV709_pythia8/[1]	0.0002305562
WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV709_pythia8/[1]	0.0001462348
ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV709_pythia8/[1]	0.0005321759
ttH_HToZZ_4LFilter_M125_13TeV_powheg2_JHUGenV709_pythia8/[1]	0.0003639351
bbH_HToZZTo4L_M125_13TeV_JHUGenV702_pythia8/[1]	0.0001339560
tqH_HToZZTo4L_M125_TuneCP5_13TeV-jhugenv7011pythia8/[1]	0.0000857830
Name of background data set	$\sigma \times \mathcal{B}$ (pb)
ZZTo4L_13TeV_powheg_pythia8/[1]	1.256
GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8/[1]	0.00158549
GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8/[1]	0.00158549
GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8/[1]	0.00158549
GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8/[1]	0.0031942
GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8/[1]	0.0031942
GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8/[1]	0.0031942

Table 5-8. Names of simulated signal and background samples for 2017 data.

[2] “RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9/MINIAODSIM”

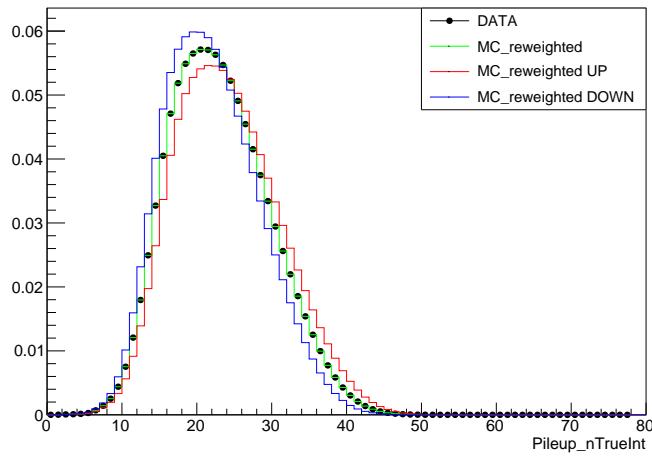
Name of signal data set	$\sigma \times \mathcal{B}$ (pb)
GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/[2]	0.01333521
VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/[2]	0.001038159
WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV7011_pythia8/[2]	0.0002305562
WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV7011_pythia8/[2]	0.0001462348
ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV7011_pythia8/[2]	0.0005321759
ttH_HToZZ_4LFilter_M125_13TeV_powheg2_JHUGenV7011_pythia8/[2]	0.0003639351
bbH_HToZZTo4L_M125_13TeV_JHUGenV7011_pythia8/[2]	0.0001339560
tqH_HToZZTo4L_M125_TuneCP5_13TeV-jhugenv7011pythia8/[1]	0.0000857830
Name of background data set	$\sigma \times \mathcal{B}$ (pb)
ZZTo4L_13TeV_powheg_pythia8/[2]	1.256
GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8/[2]	0.00158549
GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8/[2]	0.00158549
GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8/[2]	0.00158549
GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8/[2]	0.0031942
GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8/[2]	0.0031942
GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8/[2]	0.0031942

Table 5-9. Names of simulated signal and background samples for 2018 data.

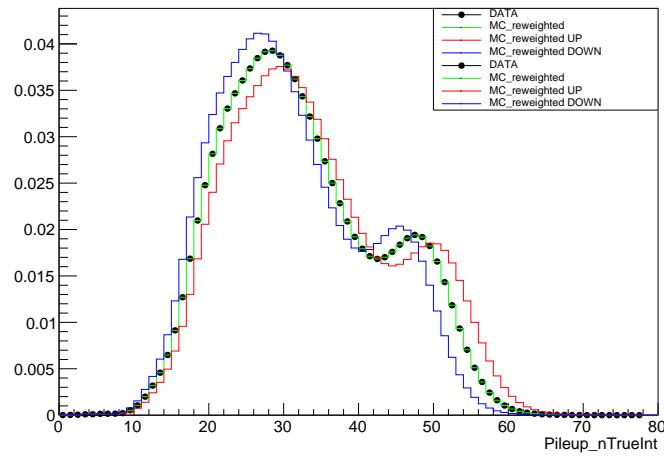
[3] “RunIISummer20UL18MiniAODv2-
106X_upgrade2018_realistic_v16_L1v1/MINIAODSIM”

Name of signal data set	$\sigma \times \mathcal{B}(\text{pb})$
GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/[3]	0.01333521
VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV7011_pythia8/[3]	0.001038159
WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV7011_pythia8/[3]	0.0002305562
WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV7011_pythia8/[3]	0.0001462348
ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV7011_pythia8/[3]	0.0005321759
ttH_HToZZ_4LFilter_M125_13TeV_powheg2_JHUGenV7011_pythia8/[3]	0.0003639351
bbH_HToZZTo4L_M125_TuneCP2_13TeV-JHUGenV7011_pythia8/[3]	0.0001339560
tqH_HToZZTo4L_M125_TuneCP5_13TeV-jhugenv7011pythia8/[1]	0.0000857830
Name of background data set	$\sigma \times \mathcal{B}(\text{pb})$
ZZTo4L_TuneCP5_13TeV_powheg_pythia8/[3]	1.256
GluGluToContinToZZTo4e_13TeV_MCFM701_pythia8/[3]	0.00158549
GluGluToContinToZZTo4mu_13TeV_MCFM701_pythia8/[3]	0.00158549
GluGluToContinToZZTo4tau_13TeV_MCFM701_pythia8/[3]	0.00158549
GluGluToContinToZZTo2e2mu_13TeV_MCFM701_pythia8/[3]	0.0031942
GluGluToContinToZZTo2e2tau_13TeV_MCFM701_pythia8/[3]	0.0031942
GluGluToContinToZZTo2mu2tau_13TeV_MCFM701_pythia8/[3]	0.0031942

A
Python unweighted



B
Python unweighted



C
Python unweighted

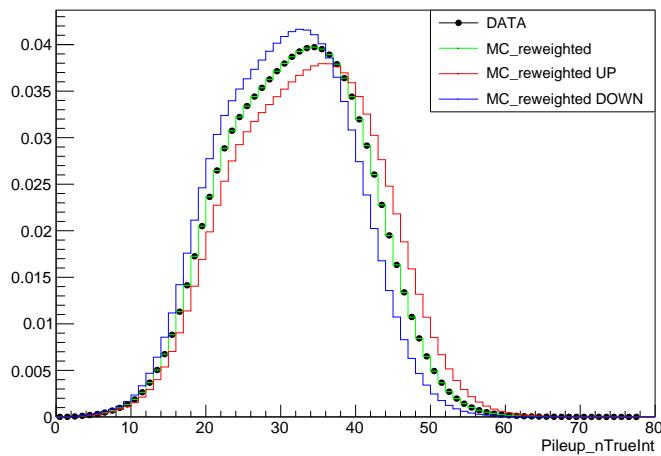


Figure 5-5. TODO:REWORD Data–simulation pileUp distributions: left 2016 (for pre and post-VFP), right 2017, bottom 2018. Up (down) scale has been obtained using 72.4 (66)mb.

5.4 Event Selection

Amidst the chaotic particle deluge that arises from pp collisions, events are carefully selected that appear to follow the decay chain and properties of the $H \rightarrow ZZ^* \rightarrow 4\ell$ signal. Crafting a well-designed *event selection* that identifies these signal events—while throwing away the non-signal events—is the goal. By implementing a rigorous trigger selection, vertex selection, final-state object selection, and ZZ-candidate selection, the purity of the signal process is optimized and, by extension, so is the precision on the measurement of m_H . The aforementioned selections are summarized in Ref. [?].

If an event contains more than one ZZ candidate that passes the selection criteria, then the candidate with the highest value of $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ is selected as the overall ZZ candidate for the event.

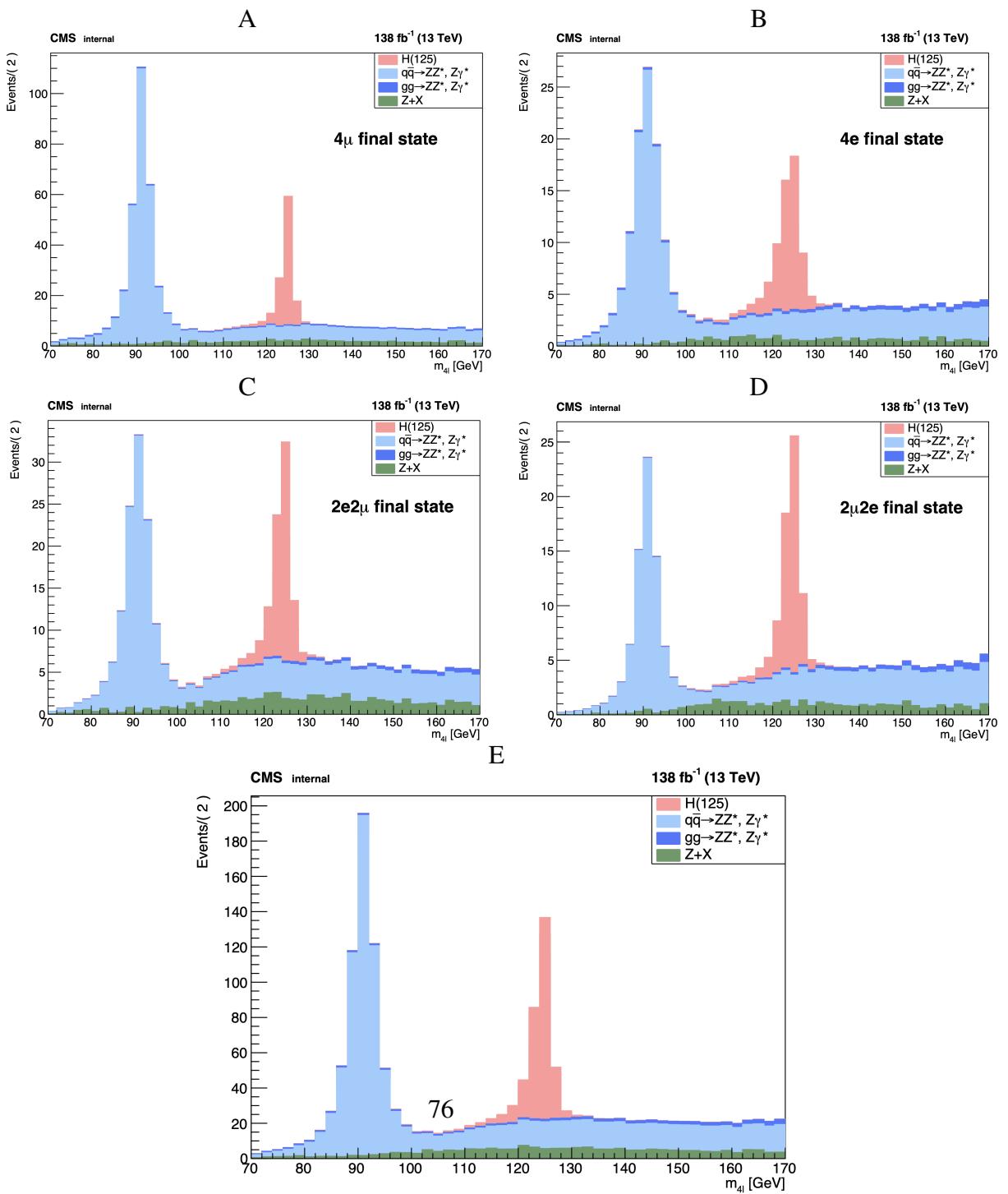
After the full analysis selection is implemented in each of the 4ℓ final states, the expected and observed yields for the signal and background processes are given in Tables 5-10 and 5-11, which count events within the narrow signal region ($105 < m_{4\ell} < 140 \text{ GeV}$) and wide signal region ($70 < m_{4\ell} < 170 \text{ GeV}$), respectively.

Table 5-10. Number of expected and observed yields of signal and background processes within the narrow signal region $105 < m_{4\ell} < 140 \text{ GeV}$ after the full event selection using 137.1 fb^{-1} , split into the four final states.

Process	4μ	$4e$	$2e2\mu$	$2\mu2e$	Inclusive
$q\bar{q} \rightarrow ZZ \rightarrow 4\ell$	88.8	38.5	63.7	41.8	232.8
$gg \rightarrow ZZ \rightarrow 4\ell$	9.7	4.8	4.8	3.7	23.0
RB	32.4	12.2	29.2	18.6	92.4
Sum of Background	130.9	55.5	97.7	64.1	348.2
Signal ($m_H = 125 \text{ GeV}$)	90.5	48.2	64.6	53.0	256.3
Total Expected	221.4	103.7	162.3	117.1	604.5
Total Observed	—	—	—	—	—

Table 5-11. Number of expected and observed yields of signal and background processes within the wide signal region $70 < m_{4\ell} < 170$ GeV after the full event selection using 137.1 fb^{-1} , split into the four final states.

Process	4μ	4e	2e2μ	2μ2e	Inclusive
$q\bar{q} \rightarrow ZZ \rightarrow 4\ell$	486.7	192.0	246.0	170.1	1094.9
$gg \rightarrow ZZ \rightarrow 4\ell$	29.7	15.2	13.3	12.2	70.5
RB	70.1	30.3	61.5	42.2	204.1
Sum of Background	586.5	237.5	320.8	224.5	1369.3
Signal ($m_H = 125$ GeV)	92.4	49.6	66.5	54.3	262.8
Total Expected	679.0	287.2	387.4	278.8	1632.3
Total Observed	—	—	—	—	—



5.5 Background Estimation

Measurement of the Higgs boson mass requires the accurate modeling of the total event yield in the signal region (SR), into which events can be categorized as either *signal* or *background*. These background events pass the signal event selection and thus spoil the purity of the signal events. This introduces further uncertainty into the final Higgs boson mass measurement. Therefore, it is a priority to reduce and to predict the expected number of background events, which can be split into two types: *irreducible background* (IB) and *reducible background* (RB) processes.

5.5.1 Irreducible Background

IB processes produce two Z bosons and each Z subsequently decays into two prompt leptons (leptons that emerge directly from the primary vertex). This reliably produces a 4ℓ final state, whose prompt leptons are typically reconstructed as four leptons that pass tight selection (*PTS leptons*), as defined in Sections ?? and ???. The IB event then looks indistinguishable from the 4 PTS lepton of the *signal* process and cannot be reduced; Thus, throwing away IB process could mean throwing away a signal event. Since IB events cannot be completely eliminated—or *reduced*—from the SR, they are given the name *irreducible* backgrounds. The two IBs for this analysis are:

- $gg \rightarrow ZZ \rightarrow 4\ell$ (gluon-gluon fusion),
- $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$ (quark-antiquark annihilation).

5.5.2 Reducible Background

While IB processes produce 4 prompt leptons, RB processes produce varying numbers of prompt and nonprompt leptons. Nonprompt leptons emerge from 3 main sources:

- misidentifying light-flavor hadrons (e.g., π^\pm) as leptons,
- heavy-flavor hadrons that decay mid-flight into leptons,
- and the asymmetric conversion of photons into electrons.

This analysis is tailored to efficiently reconstruct prompt(nonprompt) leptons as PTS(FTS) leptons. Ideally, the SR would contain only the events with 4 prompt leptons, which would always be tagged as 4 PTS leptons. However, sometimes a truly nonprompt lepton is misidentified as a PTS lepton (sometimes called a *fake lepton*), depending on the kinematic properties of the lepton. This misidentification rate (f , sometimes called the *fake rate*) is due to imperfect detector performance, inefficiencies in reconstruction, and the specific choice of lepton selections used in the analysis.

If an event produces prompt and nonprompt leptons but is reconstructed as 4 PTS leptons (a 4P event), then it contaminates the SR. These processes constitute the RB processes (sometimes called “Z + X”). Examples of RB processes for this analysis include:

- Z + jets (yields 2 prompt leptons)
- $t\bar{t}$ + jets (yields 2 prompt leptons)
- WZ + jets (yields 3 prompt leptons)
- ZZ/Z γ^* + jets (yields 4 prompt leptons).

The careful estimation of these RB contributions to the 4P region is necessary for the precise measurement of the Higgs boson mass.

5.5.2.1 OS Method

The goal of the OS Method is to estimate the number of opposite-sign same-flavor (OSSF) 4ℓ events produced by RB process that “contaminate” the 4P region (N_{4P}^{RB}), given by Eq. 5-13. However, the typical approach of using simulated samples does not model RB well, since RB processes (e.g., Z + jets) rely on higher-order effects like jet modelling which are not yet accurately simulated. Instead, a data-driven approach is used.

The logic of the OS Method is to study events in data that are similar to, but not exactly the same as, those found in the 4P region. Thus, events in data are sorted into 2 control regions (CRs), both of which are orthogonal to the 4P region and to each other:

- the 3P1F CR (built from 3 PTS and 1 FTS leptons)
- the 2P2F CR (built from 2 PTS and 2 FTS leptons).

The event selection for the 2P2F and 3P1F CRs is almost identical to that of the SR, except that the FTS lepton(s) are required to build the Z_2 . The events that contribute mostly to the 2P2F(3P1F) CR are those that produce 2(3) prompt and 2(1) nonprompt leptons and are called $2pr(3pr)$ events. Similarly, the events that contribute mostly to the 4P SR are those that produce 4 prompt leptons and are called $4pr$ events.

The final formula for N_{4P}^{RB} is obtained by first supposing that event k is a 2pr event and contributes to the 2P2F CR an event weight of $w_{2pr \rightarrow 2P2F}^k$. This weight is built from \hat{w}^k , which is the product of analysis weights (pileup, L1 pre-firing, etc.), and the reconstruction efficiencies (ϵ) of each lepton (ℓ_n):

$$w_{2pr \rightarrow 2P2F}^k = \hat{w}^k \cdot \epsilon_P^{pr}(\ell_1^k) \cdot \epsilon_P^{pr}(\ell_2^k) \cdot \epsilon_F^{np}(\ell_3^k) \cdot \epsilon_F^{np}(\ell_4^k), \quad (5-1)$$

where the superscript of ϵ indicates the lepton promptness (pr = prompt, np = nonprompt), the subscript indicates the lepton tightness status (P = PTS, F = FTS). To simplify the equations that follow, \hat{w}^k is set to unity. If the reconstruction efficiencies of a particular category are the same for all j leptons across all k events (e.g., $\epsilon_P^{pr}(\ell_j^k) \equiv \epsilon_P^{pr}$), then Eq. 5-1 reduces to:

$$w_{2pr \rightarrow 2P2F}^k = \left(\epsilon_P^{pr} \right)^2 \left(\epsilon_F^{np} \right)^2. \quad (5-2)$$

Although a 2pr event mostly contributes to the 2P2F CR, a nonprompt lepton may be misidentified as a PTS lepton, depending on ϵ_P^{np} . Such an event would then fall into the 3P1F CR and, allowing for only one nonprompt PTS lepton at a time, contributes an effective weight of

$$\begin{aligned} w_{2pr \rightarrow 3P1F}^k &= \left(\epsilon_P^{pr} \right)^2 \left[\epsilon_P^{np}(\ell_1^k) \cdot \epsilon_F^{np}(\ell_2^k) + \epsilon_F^{np}(\ell_1^k) \cdot \epsilon_P^{np}(\ell_2^k) \right] \\ &= \left(\epsilon_P^{pr} \right)^2 [2\epsilon_P^{np}\epsilon_F^{np}]. \end{aligned} \quad (5-3)$$

Using the fact that a (non)prompt lepton is exclusively either PTS or FTS ($\epsilon_P^{pr(np)} + \epsilon_F^{pr(np)} = 1$), while recognizing that $\epsilon_P^{np} \equiv f$ (Sec. 5.5.2) and defining $\epsilon_P^{pr} \equiv \epsilon$, allows Eq. 5-3 to be written as

$$w_{2pr \rightarrow 3P1F}^k = 2\epsilon^2 f(1-f). \quad (5-4)$$

Even more rarely, both nonprompt leptons from a 2pr event may be misidentified as PTS leptons. In this case, event k contributes to the 4P region an effective weight of

$$w_{2\text{pr} \rightarrow 4\text{P}}^k = \epsilon^2 f^2, \quad (5-5)$$

where it is assumed that both leptons have the same misidentification rate. Similar equations can be derived for the contributions of a 3pr event to the 3P1F CR and to the 4P region:

$$w_{3\text{pr} \rightarrow 3\text{P1F}}^k = \epsilon^3 (1 - f) \quad (5-6)$$

$$w_{3\text{pr} \rightarrow 4\text{P}}^k = \epsilon^3 f. \quad (5-7)$$

Since a 3pr event needs only 1 nonprompt PTS lepton to be included in the 4P region (therefore, carrying only 1 factor of f), a 3pr event tends to contribute more weight to 4P than does a 2pr event (which carries f^2).

If the total number of 2pr, 3pr, and 4pr events is $X_{2\text{pr}}$, $X_{3\text{pr}}$, and $X_{4\text{pr}}$, respectively, then Fig. 5-7 shows how the weight of a single event, derived in the previous equations (and others forthcoming), from each category contributes to the final yield of each CR ($N_{2\text{P2F}}$, $N_{3\text{P1F}}$) and of the SR ($N_{4\text{P}}$). It is then straightforward to evaluate the expected number of RB 4P events:

$$\begin{aligned} N_{4\text{P}}^{\text{RB}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 4\text{P}}^k + \sum_{m=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 4\text{P}}^m \\ &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 f^2 + \sum_{m=1}^{X_{3\text{pr}}} \epsilon^3 f \\ &= f^2 \epsilon^2 X_{2\text{pr}} + f \epsilon^3 X_{3\text{pr}}, \end{aligned} \quad (5-8)$$

where only the quantities f and ϵ are known, so $X_{2\text{pr}}$ and $X_{3\text{pr}}$ must be estimated. This is achieved

by relating $X_{2\text{pr}}$ to $N_{2\text{P}2\text{F}}$ using Eq. 5-2 across all 2pr events:

$$\begin{aligned}
 N_{2\text{P}2\text{F}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 2\text{P}2\text{F}}^k \\
 &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 (1-f)^2 \\
 &= (1-f)^2 \epsilon^2 X_{2\text{pr}}
 \end{aligned} \tag{5-9}$$

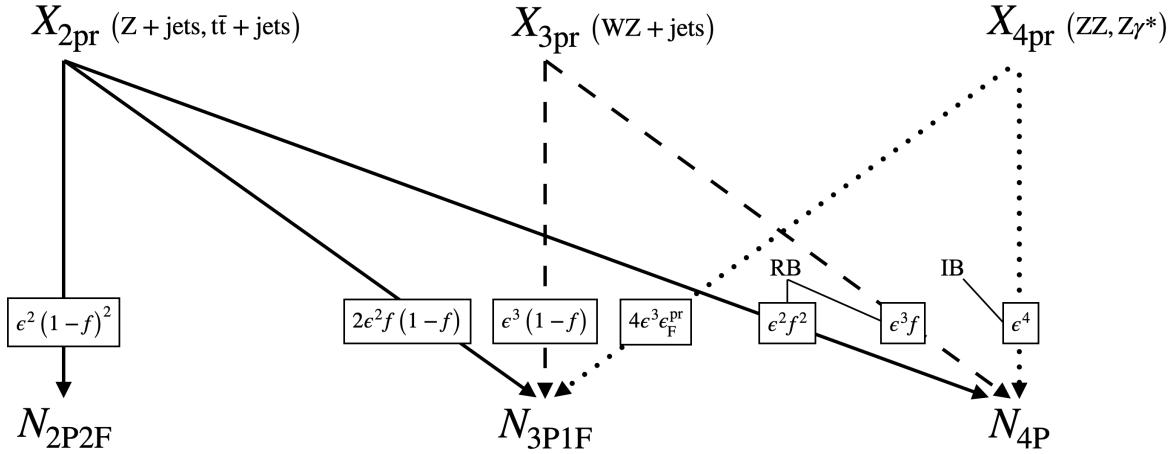


Figure 5-7. Contributions of per-event weights (boxed values) of various n -prompt-lepton processes ($X_{n\text{pr}}$, in parentheses) to the total numbers of events in the observed control regions ($N_{2\text{P}2\text{F}}, N_{3\text{P}1\text{F}}$) and signal region ($N_{4\text{P}}$). The labels “RB” and “IB” indicate those contributions which comprise the reducible and irreducible backgrounds, respectively.

The strategy to relate $X_{3\text{pr}}$ to $N_{3\text{P}1\text{F}}$ is not as straightforward as relating $X_{2\text{pr}}$ to $N_{2\text{P}2\text{F}}$, since two other sources also contribute to the 3P1F CR (as shown in Fig. 5-7):

- A 2pr RB process can yield one nonprompt PTS lepton (via Eq. 5-4).
- A 4-prompt IB process ($q\bar{q}/gg \rightarrow ZZ \rightarrow 4\ell$, “ZZ”) can yield one FTS lepton.

The second of these is well estimated from simulation, since ZZ produces 4 prompt leptons. If the total number of simulated ZZ events is $X_{4\text{pr}}^{\text{ZZ}}$, then event k belonging to these events contributes to the 3P1F CR an effective weight of

$$w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P}1\text{F}}^k = 4\epsilon^3 \epsilon_F^{\text{pr}}, \tag{5-10}$$

which accounts for any of the 4 prompt leptons to be reconstructed as a FTS lepton while the others are PTS leptons. Incorporating Eqs. 5-4, 5-6, and 5-10 for all events relates $X_{3\text{pr}}$ to $N_{3\text{P1F}}$:

$$\begin{aligned}
 N_{3\text{P1F}} &= \sum_{j=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 3\text{P1F}}^j + \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 3\text{P1F}}^k + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P1F}}^m \\
 &= \sum_{j=1}^{X_{3\text{pr}}} \epsilon^3 (1-f) + \sum_{k=1}^{X_{2\text{pr}}} 2\epsilon^2 f (1-f) + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} 4\epsilon^3 \epsilon_{\text{F}}^{\text{pr}} \\
 &= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + 4\epsilon^3 \epsilon_{\text{F}}^{\text{pr}} X_{4\text{pr}}^{\text{ZZ}} \\
 &= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + N_{3\text{P1F}}^{\text{ZZ}}, \tag{5-11}
 \end{aligned}$$

where $N_{3\text{P1F}}^{\text{ZZ}}$ is simply the raw (integer) number of ZZ events that pass 3P1F selections, obtained directly from simulation.

At this point, $N_{4\text{P}}^{\text{RB}}$ can be isolated by combining Eqs. 5-8, 5-9, and 5-11:

$$N_{4\text{P}}^{\text{RB}} = \left(\frac{f}{1-f} \right) N_{3\text{P1F}}^{\text{Data}} - \left(\frac{f}{1-f} \right)^2 N_{2\text{P2F}}^{\text{Data}} - \left(\frac{f}{1-f} \right) N_{3\text{P1F}}^{\text{ZZ}}. \tag{5-12}$$

Thus, Eq. 5-12 estimates the RB contribution to the 4P region by using a single lepton misidentification rate to reweight the raw yields of events found in the 3P1F and 2P2F CRs of data, and also the 3P1F CR of ZZ.

It should be mentioned that the above formula was derived assuming that the per event analysis weights were set to unity ($\hat{w}^k = 1$) before being scaled by the misidentification rates. It was also assumed that f is a constant, for all nonprompt leptons misidentified as PTS leptons across all events, which is not the case as is shown in Figures 5-16 and 5-17. Thus, an extension of Eq. 5-12 can be formed by assigning misidentification rates that depend on the kinematical variables per lepton, by restoring the analysis weights per event, and by summing over the total number of raw yields per CR:

$$N_{4\text{P}}^{\text{RB}} = \sum_{i=1}^{N_{3\text{P1F}}^{\text{Data}}} \hat{w}_i^i \frac{f_i}{1-f_i} - \sum_{j=1}^{N_{2\text{P2F}}^{\text{Data}}} \hat{w}_j^j \frac{f_{1,j}}{1-f_{1,j}} \frac{f_{2,j}}{1-f_{2,j}} - \sum_{k=1}^{N_{3\text{P1F}}^{\text{ZZ}}} \hat{w}_{\text{ZZ}}^k \frac{f_k}{1-f_k}, \tag{5-13}$$

where $f_{1,j}$ and $f_{2,j}$ are the misidentification rates of the first and second FTS leptons, respectively,

found in the j^{th} 2P2F event and \hat{w}_{ZZ}^k accounts for the differential QCD and electroweak K factors.

The prediction of Eq. 5-4 can be seen in Figs. 5-12–5-15 (red line) for 2016–2018 UL data.

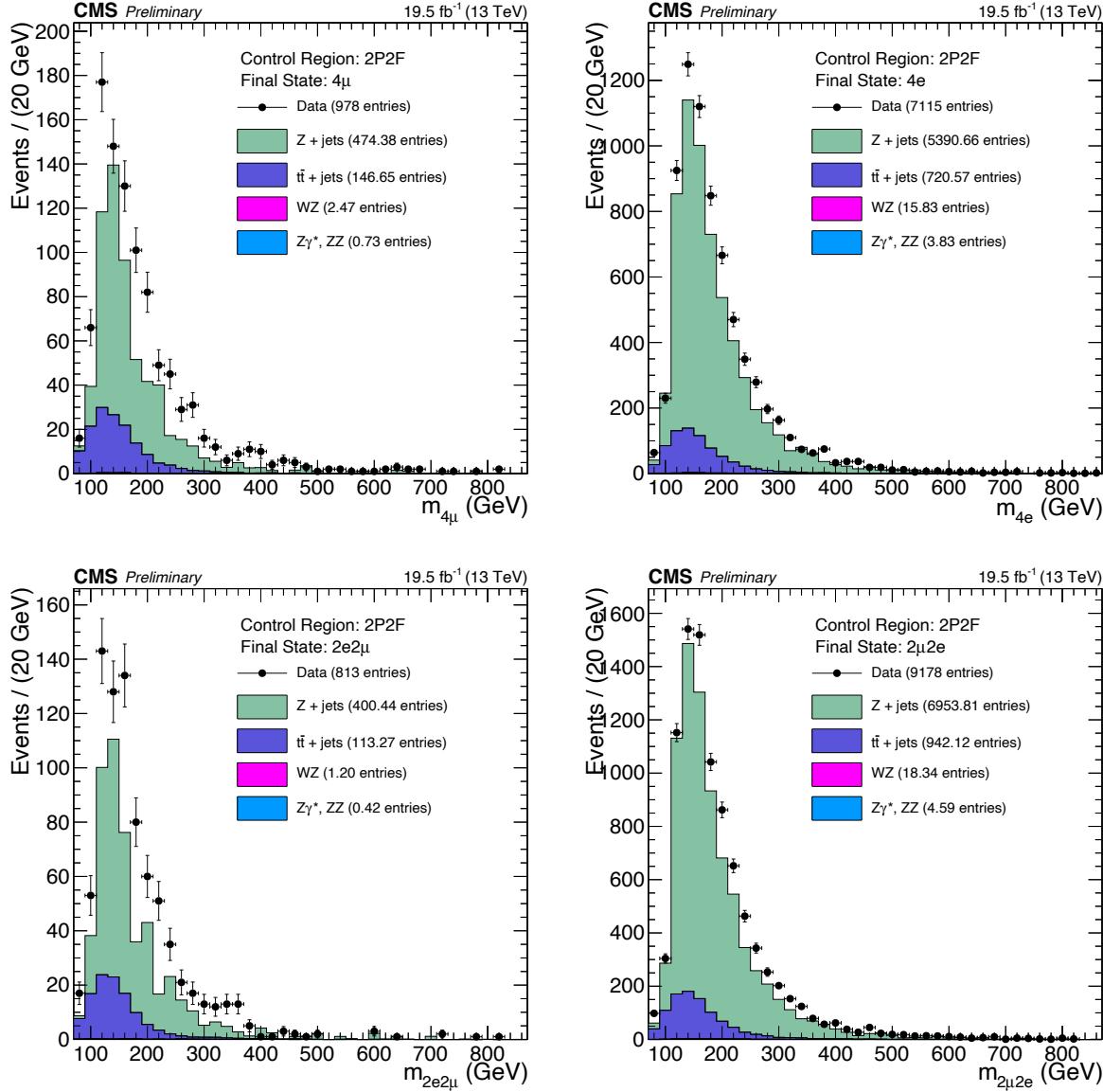


Figure 5-8. Events from 2016 pre-VFP UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , ZZ , $Z\gamma^*$). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

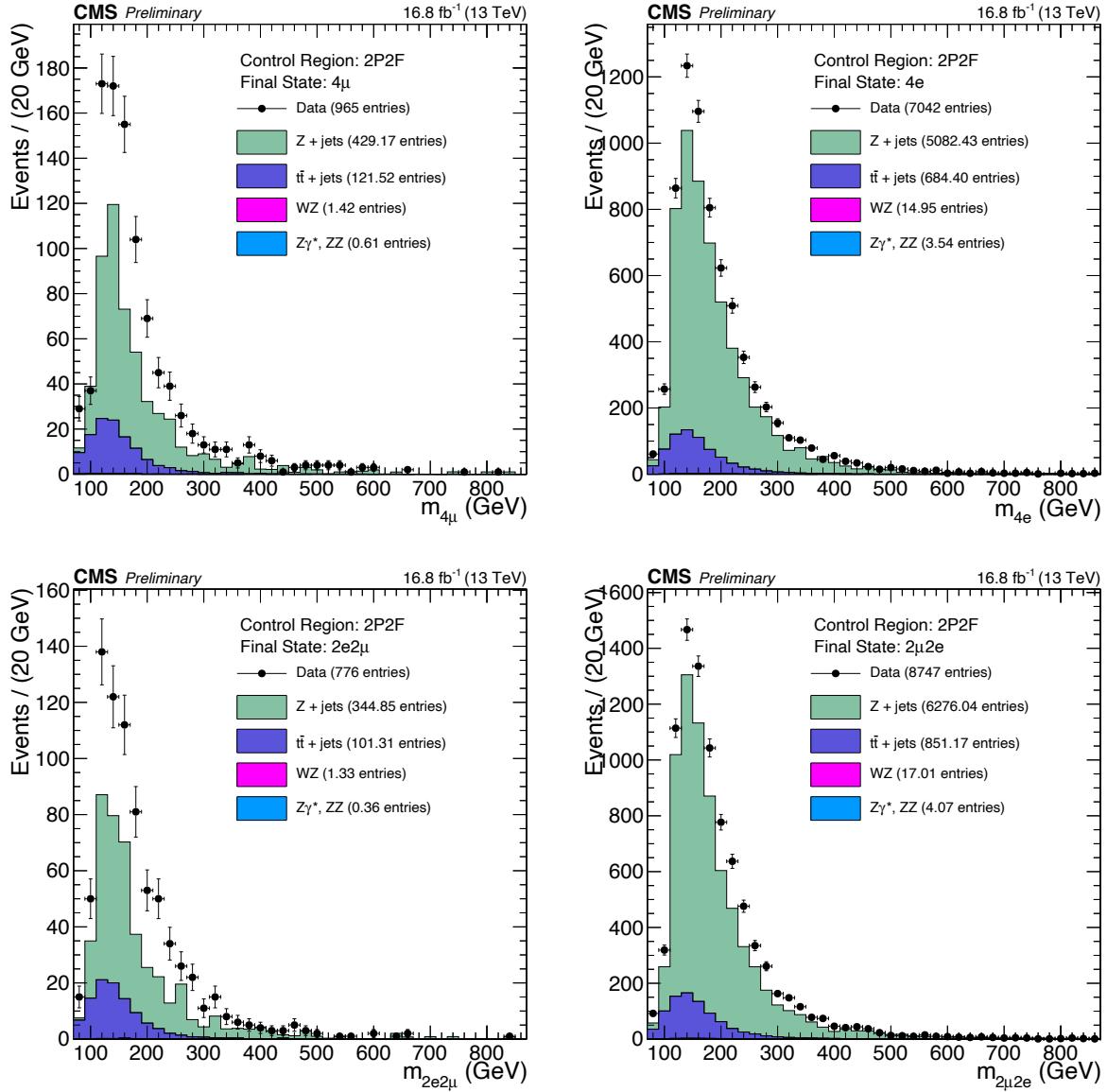


Figure 5-9. Events from 2016 post-VFP UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, t̄t + jets, WZ, ZZ, Z γ^*). The results are split into the 4 ℓ final states: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

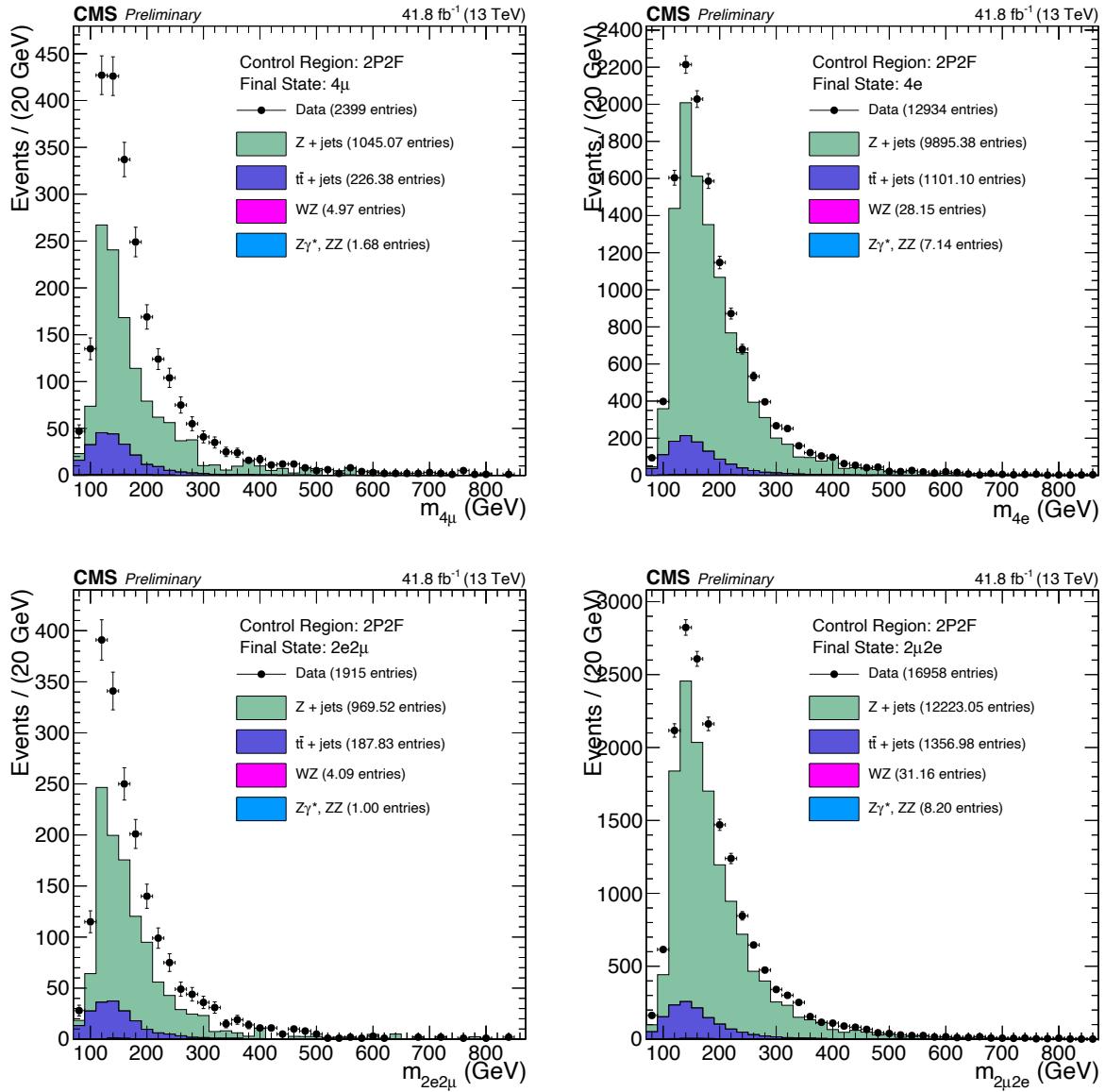


Figure 5-10. Events from 2017 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, tt + jets, WZ, ZZ, Z γ^*). The results are split into the 4 ℓ final states: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

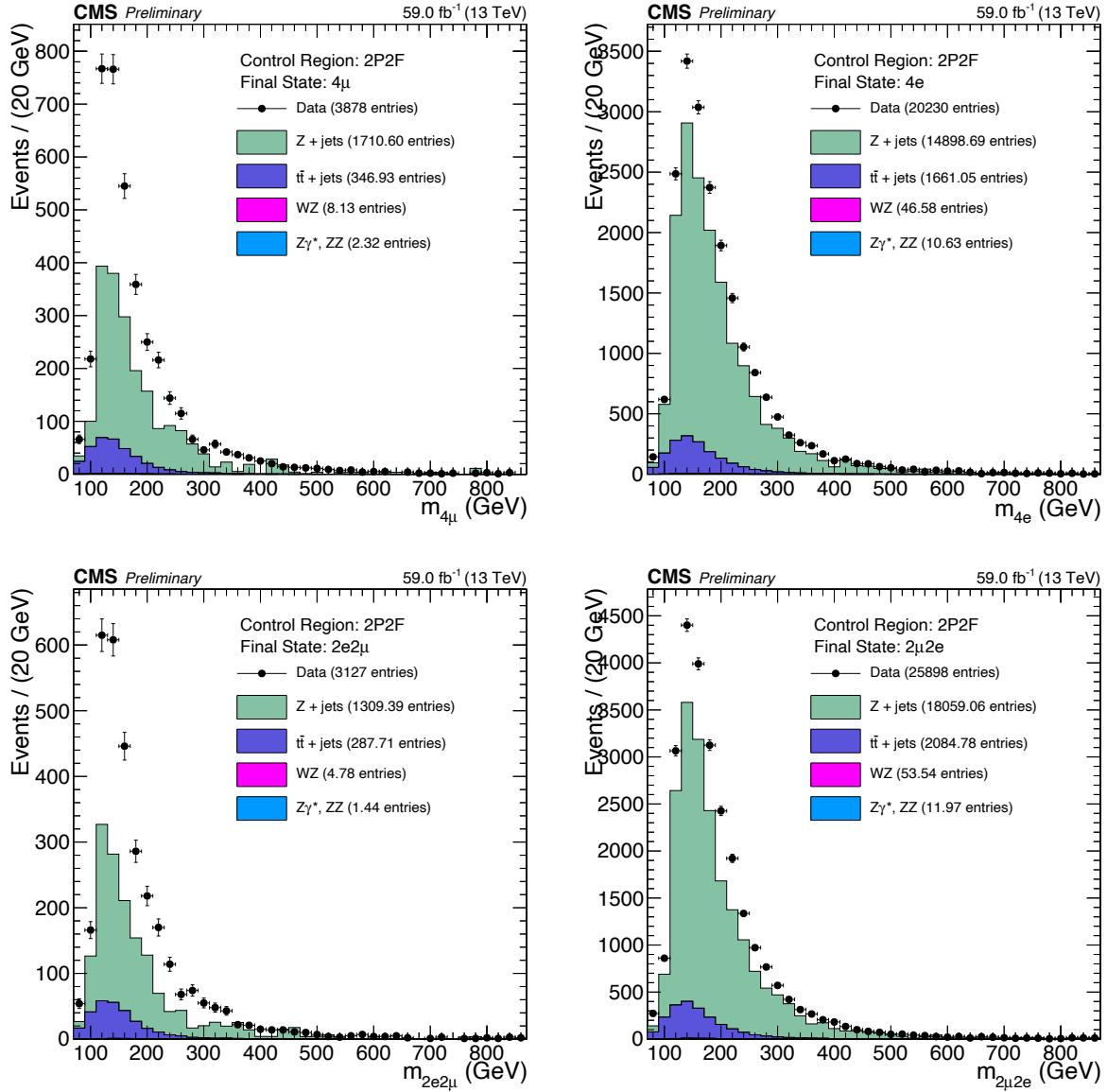


Figure 5-11. Events from 2018 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , ZZ , $Z\gamma^*$). The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

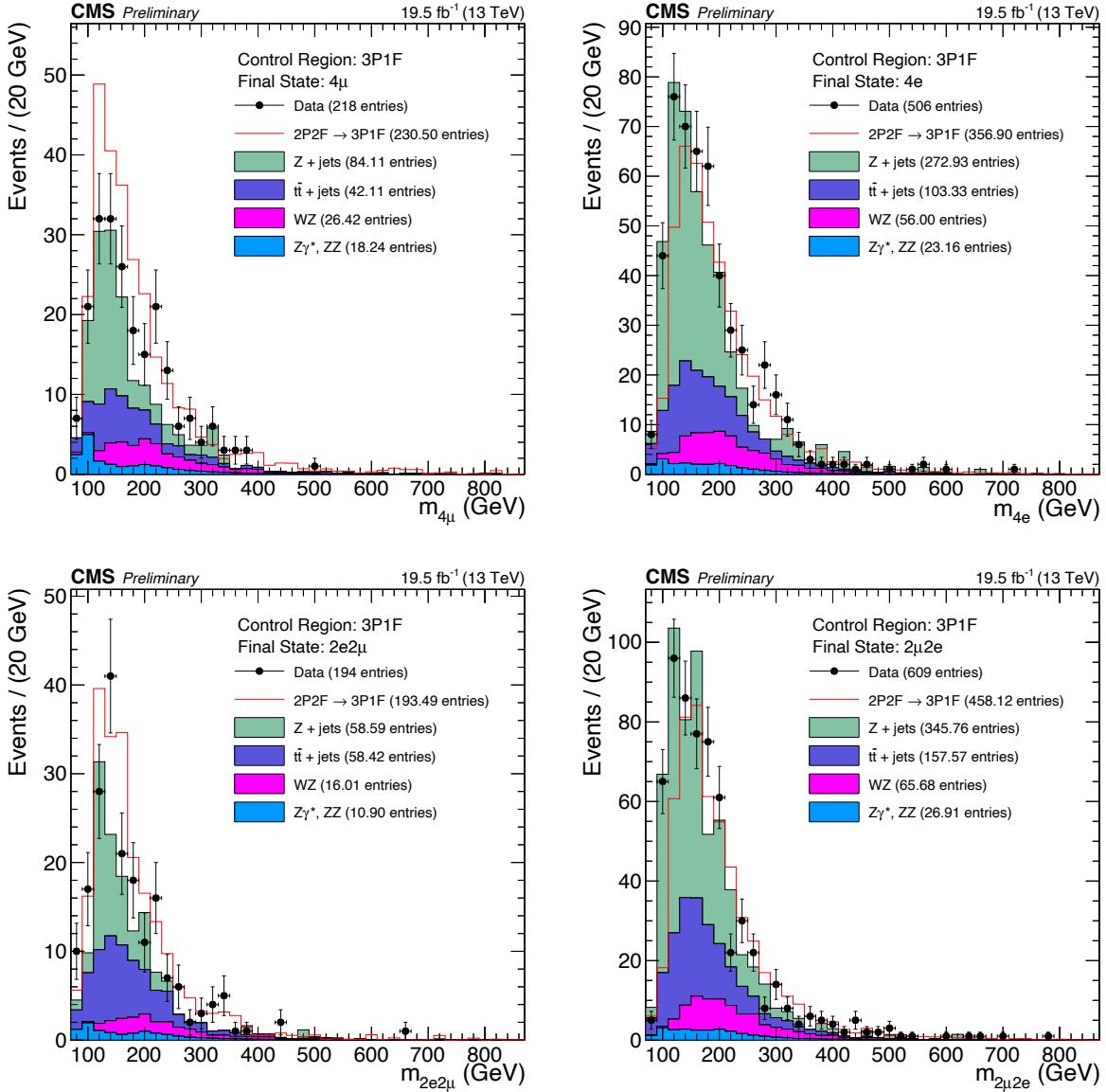


Figure 5-12. Events from 2016 pre-VFP UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, t \bar{t} + jets, WZ, ZZ, Z γ^*) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 5-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4 ℓ final states: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

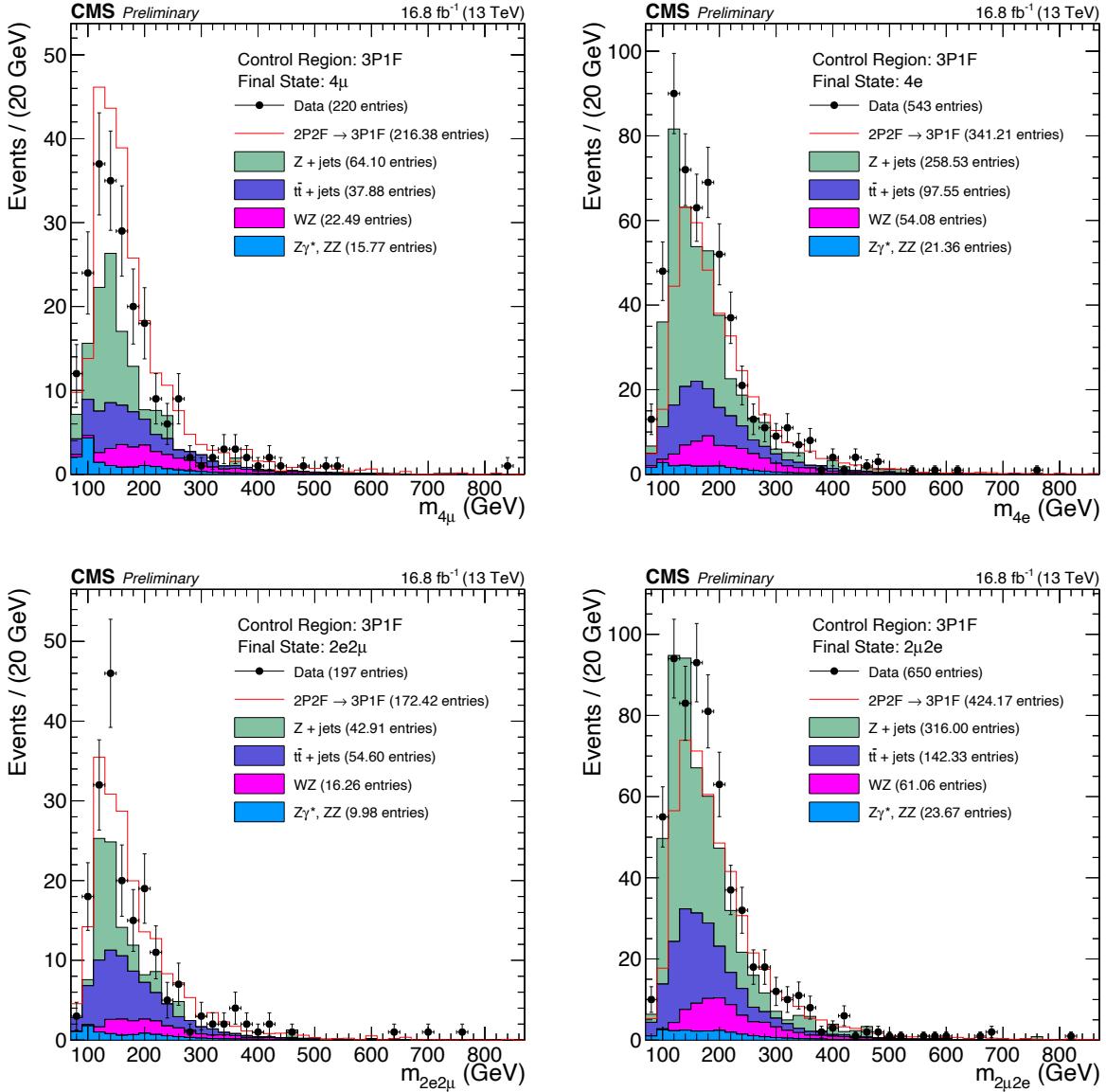


Figure 5-13. Events from 2016 post-VFP UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , ZZ , $Z\gamma^*$) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 5-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

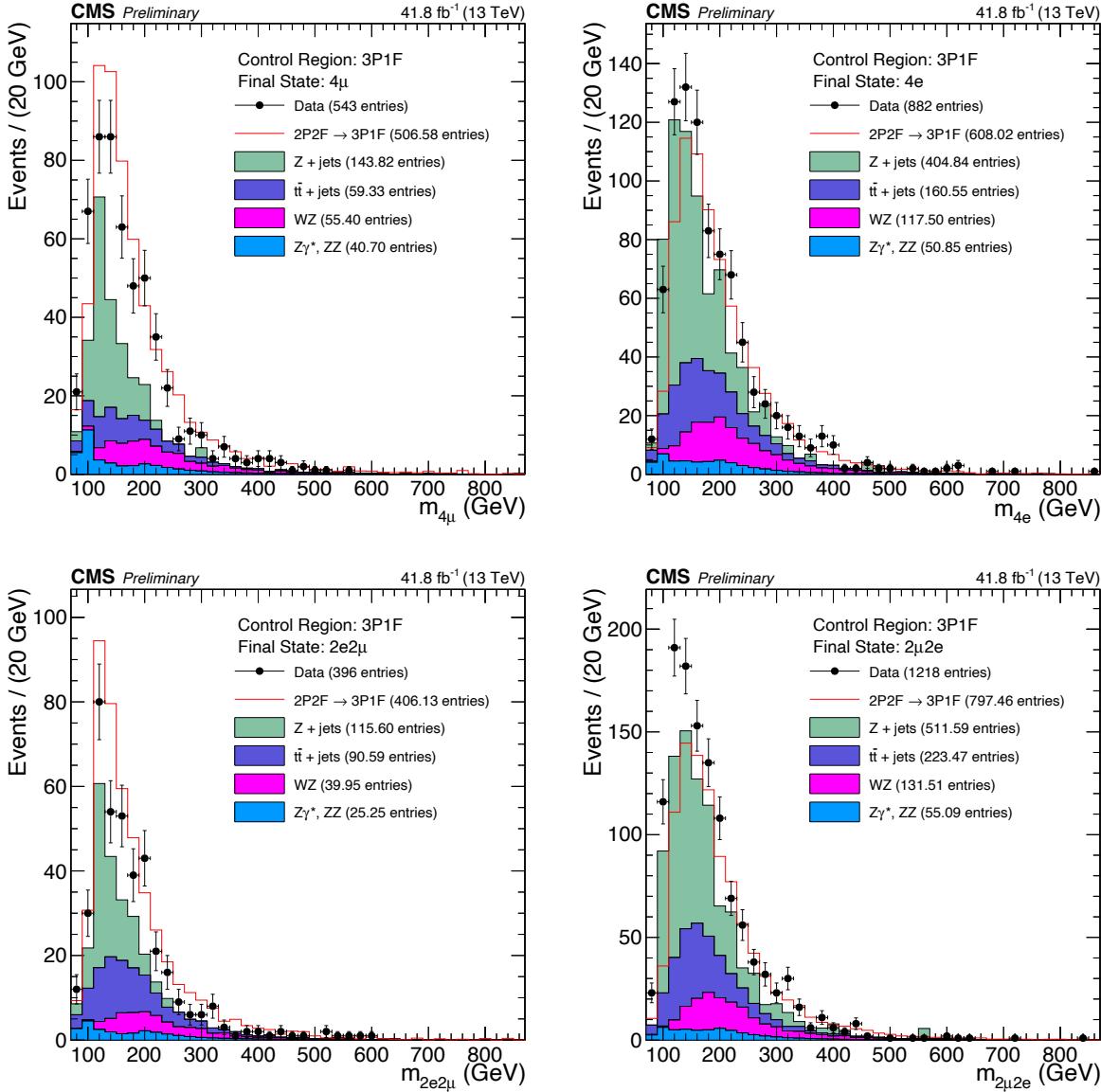


Figure 5-14. Events from 2017 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ, ZZ, $Z\gamma^*$) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 5-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

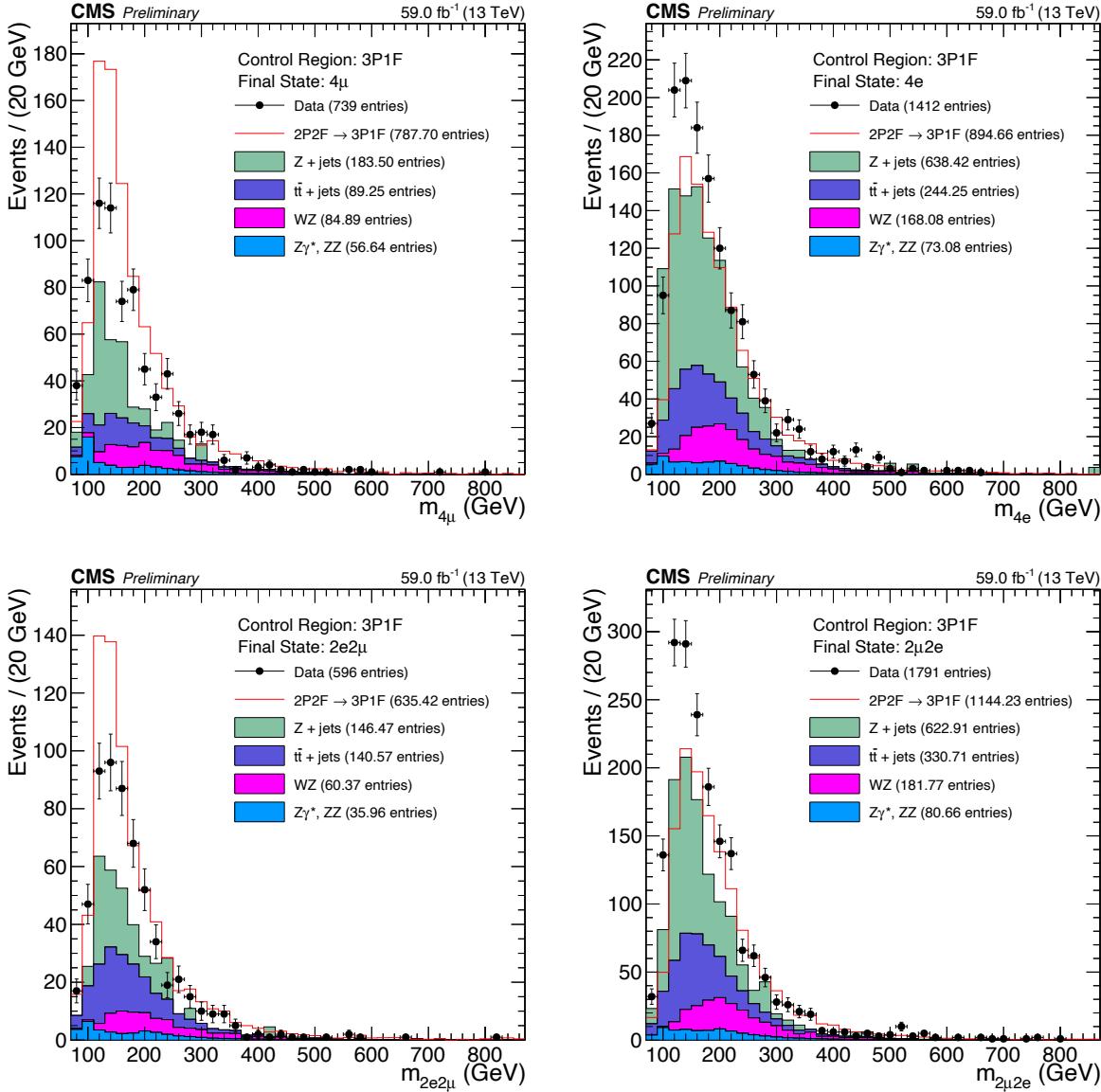


Figure 5-15. Events from 2018 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ($Z + \text{jets}$, $t\bar{t} + \text{jets}$, WZ , ZZ , $Z\gamma^*$) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 5-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), $4e$ (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

5.5.2.2 Lepton misidentification rate measurement

As mentioned in Sec. 5.5.2, the lepton misidentification rate (f) is the probability that a nonprompt lepton will pass tight selections (PTS). The value f is a function of the flavor ($\ell = e, \mu$), p_T , and η of a lepton. The misidentification rate is calculated by simply counting the number of nonprompt PTS leptons (N_P^{np}) that enter a particular $\ell, p_T, |\eta|$ bin compared to the total number of loose probe leptons (N_L^{np}) in the same bin:

$$f(\ell, |\eta|, p_T) = \frac{N_P^{\text{np}}}{N_L^{\text{np}}}. \quad (5-14)$$

The p_T^e bin edges are [5–10–20–30–40–50–80] GeV and the p_T^μ bin edges are [5–7–10–20–30–40–50–80] GeV. The nonprompt leptons used to measure f are taken from events in data with a signature like that of $Z + \ell_L$, where Z is a Z boson and ℓ_L is a loose lepton (e, μ). By construction, this region of events is orthogonal to the 2P2F, 3P1F, and 4P regions, and provides a clean source of ℓ_L . The loose lepton, whose selection is defined in Sections ?? and ??, is also called the *probe* lepton. The probe lepton is either a PTS or FTS lepton and is counted towards both the numerator and denominator of Eq. 5-14.

Events are selected that satisfy the following criteria:

- The event has exactly 3 leptons.
- The event contains $E_T^{\text{miss}} < 25$ GeV.
- Two of the leptons form a Z candidate. A Z candidate is formed when:
 - The lepton pair is OSSF.
 - Both leptons PTS.
 - The leading lepton has $p_T > 20$ GeV.
 - The subleading lepton has $p_T > 10$ GeV.
 - The lepton pair satisfies $|m_{\ell\ell} - m_{Z_{\text{PDG}}}| < 7$ GeV.
- The third and final lepton is loose (i.e., either a PTS or FTS lepton).
- Suppress QCD processes: probe lepton and OS lepton from Z have $m_{\ell\ell} > 4$ GeV.

The calculation of f requires that ℓ_L is a nonprompt lepton but since data events were used, this may not be the case. For example, the decay of WZ produces 3 prompt leptons and so this

contribution must be subtracted. Thus, the number of expected prompt probe leptons from WZ events is subtracted from both the numerator and denominator in Eq. 5-14 for each $\ell, p_T, |\eta|$ bin. The final misidentification rates used in the OS Method for electrons and muons are shown in Figures 5-16 and 5-17 using 2016–2018 UL data.

5.6 Techniques to Improve Mass Resolution

5.6.1 Event-by-event mass uncertainty

TODO:REWORD Individual lepton uncertainty on momentum measurement can be predicted on a per-lepton basis. In the case of muons, the full error matrix is obtained using muon track fit; for the electrons, instead, the momentum error is estimated from the combination of the ECAL and tracker measurement, neglecting the uncertainty on the track direction from the GSF fit.

The uncertainty on the kinematics at the per-lepton level is then propagated to the four-lepton case to predict the mass error on an event-by-event basis, using the following approach.

Each δm_i , corresponding to individual lepton momentum variation, is calculated separately and then the measured resolution on the invariant mass of the four leptons is taken as the quadrature sum of the four individual δm_i :

$$m_0 = F(p_{T1}, \phi_1, \eta_1; p_{T2}, \phi_2, \eta_2; p_{T3}, \phi_3, \eta_3; p_{T4}, \phi_4, \eta_4)$$

$$\delta m_i = F(\dots; p_{Ti} + \delta p_{Ti}, \phi_i, \eta_i; \dots) - m_0$$

$$\delta m = \sqrt{\delta m_1^2 + \delta m_2^2 + \delta m_3^2 + \delta m_4^2}$$

TODO:CONTAINS DATA. The full error matrices ($\delta p_T/p_T, \eta$) for muons and electrons, separately, are shown in Fig. 5-18 for all years.

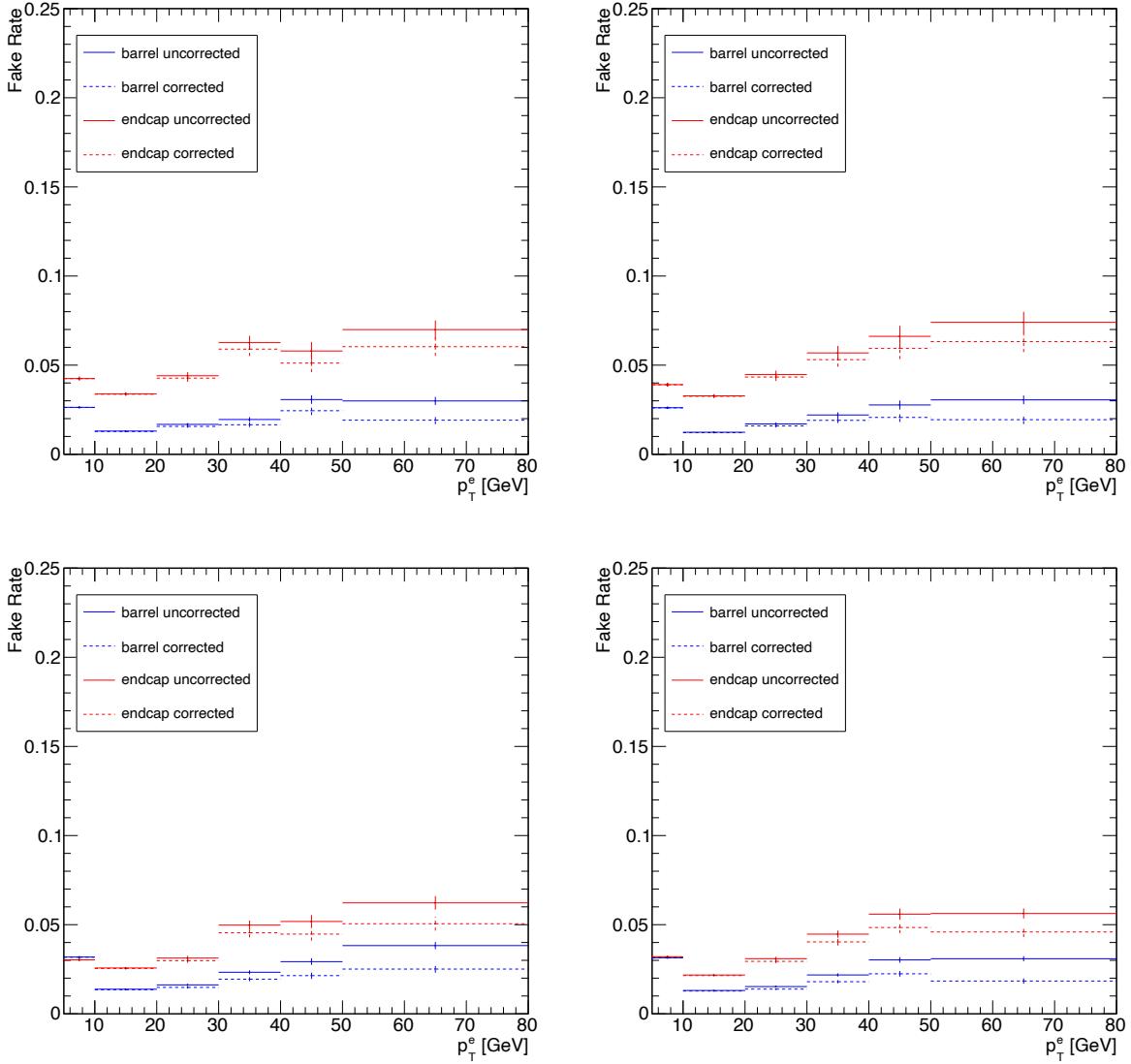


Figure 5-16. Electron misidentification rates vs. the p_T of the probe electron evaluated for the OS Method selecting $Z + e$ events in UL data for each year in Run 2: 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right). Rates are evaluated separately for the barrel (blue lines) and endcap (red lines) regions, partitioned at $|\eta^e| = 1.497$. The final rates (dashed lines) are calculated by subtracting out the expected WZ contribution—estimated from simulation—from the numerator and denominator in the rate calculation.

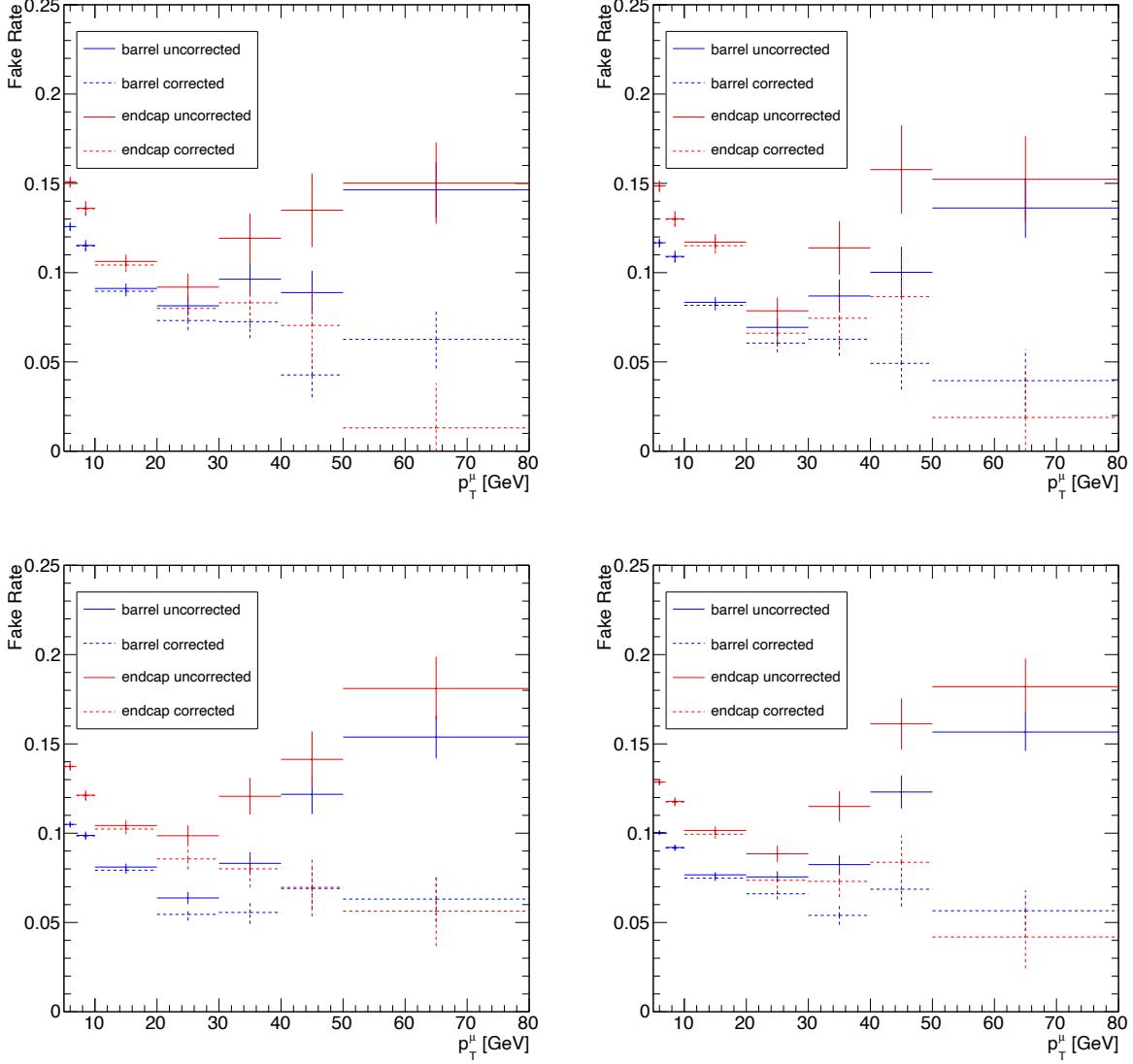


Figure 5-17. Muon misidentification rates vs. the p_T of the probe muon evaluated for the OS Method selecting $Z + \mu$ events in UL data for each year in Run 2: 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right). Rates are evaluated separately for the barrel (blue lines) and endcap (red lines) regions, partitioned at $|\eta^\mu| = 1.2$. The final rates (dashed lines) are calculated by subtracting out the expected WZ contribution—estimated from simulation—from the numerator and denominator in the rate calculation.

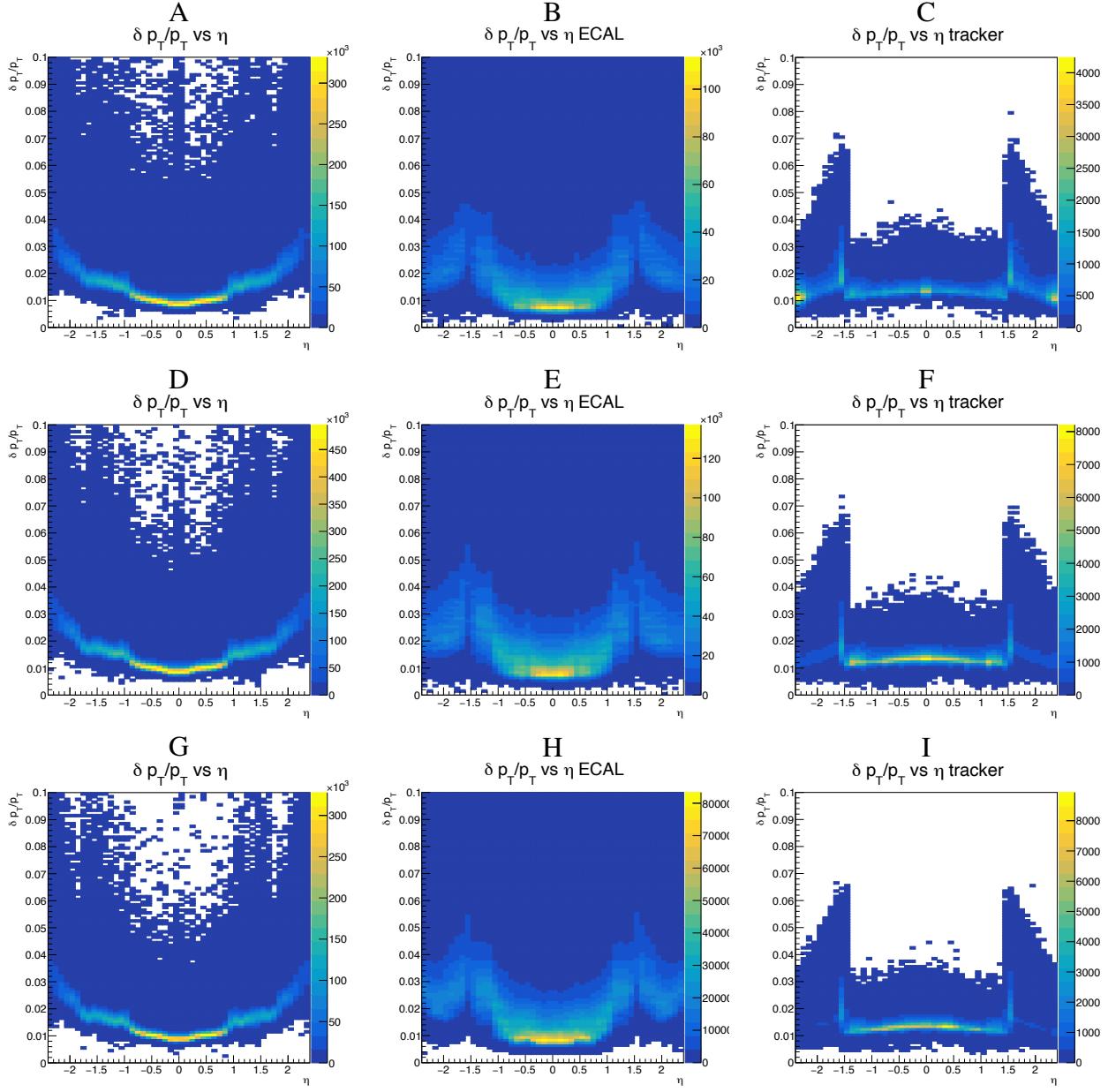


Figure 5-18. Scatter plot of the relative lepton p_T error vs η for muons (left column), ECAL driven electrons (middle column), and tracker driven electrons (right column) for 2016 (top row), 2017 (middle row), and 2018 (bottom row) data.

TODO:REWORD, CONTAINS DATA Starting from these distributions, corrections to

momentum uncertainty in mutual $|\eta|$ bins are derived for muons (Fig. 5-19), ECAL-driven electrons (Fig. 5-20), and tracker electrons (Fig. 5-21) using bins of $\delta p_T/p_T$ vs. $|\eta|$. The scatter plots $\delta p_T/p_T$ vs. p_T are shown in Figs. 5-20 and 5-21.

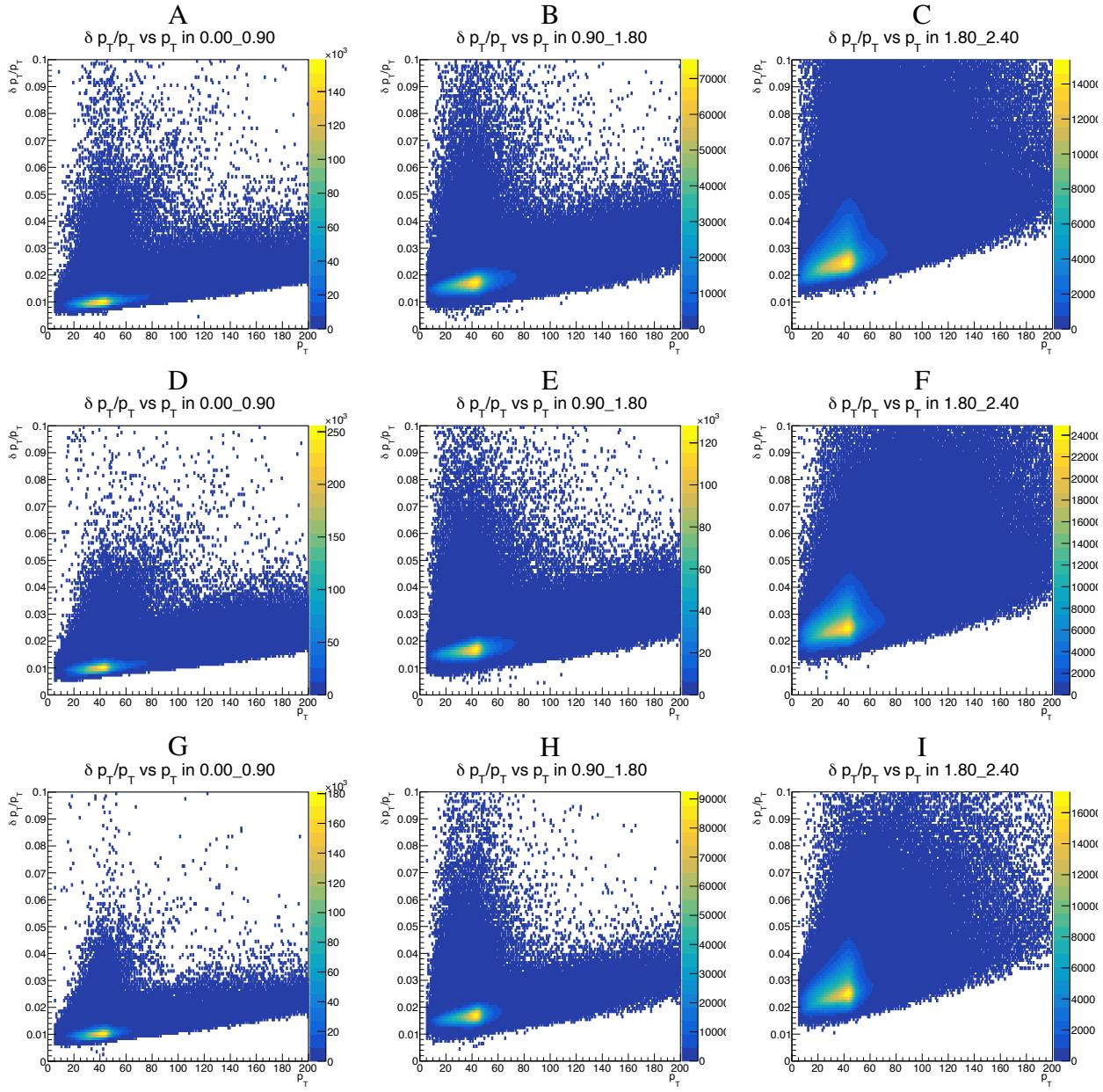


Figure 5-19. Scatter plot of the relative lepton p_T error vs p_T for muons with $0 < |\eta| < 0.9$ (left column), $0.9 < |\eta| < 1.8$ (middle column), and $1.8 < |\eta| < 2.4$ (right column) for 2016 (top row), 2017 (middle row), and 2018 (bottom row) data.

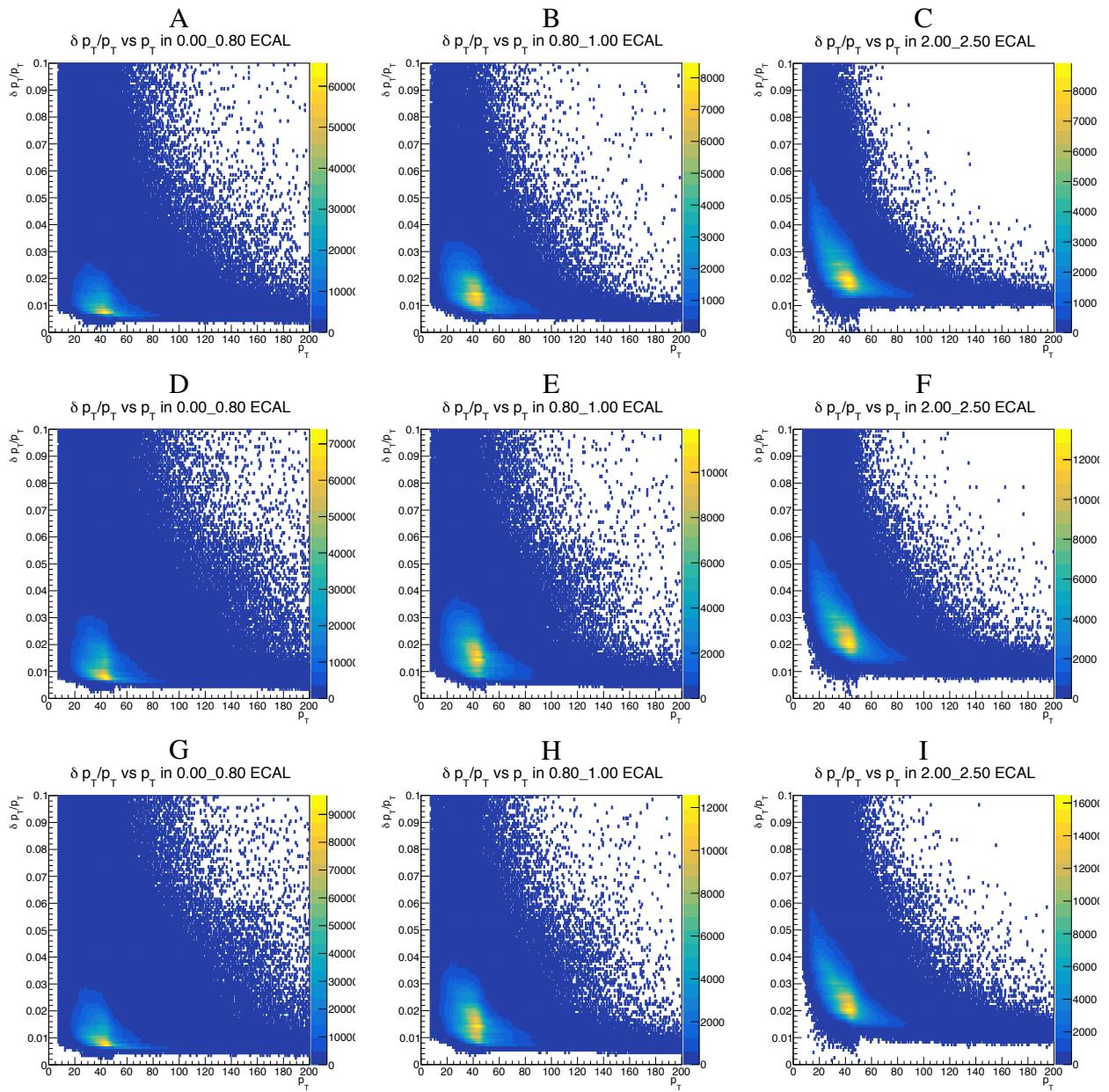


Figure 5-20. Scatter plot of the relative lepton p_T error vs p_T for ECAL-driven electrons with $0 < |\eta| < 0.8$ (left column), $0.8 < |\eta| < 1.0$ (middle column), and $2.0 < |\eta| < 2.5$ (right column) for 2016 (top row), 2017 (middle row), and 2018 (bottom row) data.

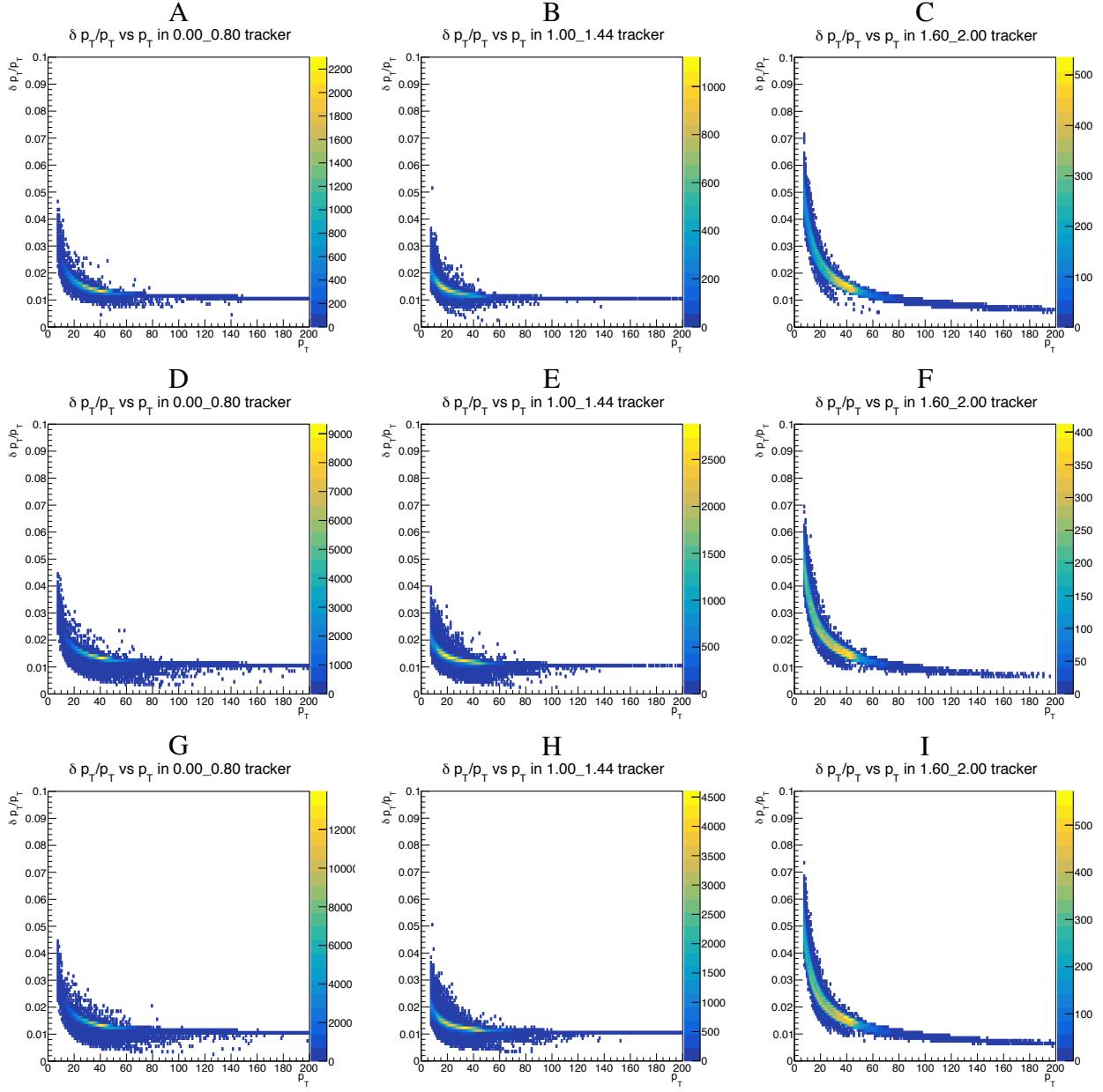


Figure 5-21. Scatter plot of the relative lepton p_T error vs p_T for tracker-driven electrons with $0 < |\eta| < 0.8$ (left column), $1.0 < |\eta| < 1.44$ (middle column), and $1.6 < |\eta| < 2.0$ (right column) for 2016 (top row), 2017 (middle row), and 2018 (bottom row) data.

5.6.2 Model and procedure to derive corrections

To derive the corrections (λ), the dilepton mass $m_{\ell\ell}$ is fitted twice with a Breit-Wigner (BW) convoluted with a Crystal Ball (CB), plus exponential function (EXP). In this model, the BW represents true m_Z shape, the CB simulates the detector effect, and the EXP describes the background. When deriving corrections, mean and sigma of Z's BW shape have been set to PDG

values ($mean_Z = 91.19$ GeV, $\sigma_Z = 2.49$ GeV [8]). The fit is done in the mass range [60, 120] GeV, using only e^+e^- or $\mu^+\mu^-$ pairs.

The first fit is used to fix all the parameters of the functions but the σ of the CB which is replaced in the second fit by $\lambda \times \delta_{m_Z}$, where λ is the floated parameter of the fit.

The summary of λ correction factors for electrons and muons is presented in Table ??.

5.6.2.1 Validation of corrections (MC, data)

A closure test is performed to validate correction derived for lepton p_T error.

First, events are divided according to different predicted $\delta m_Z/m_Z$ ranges before corection. Then, in each bin, the dilepton mass distribution is fitted using a BW convoluted with CB plus exponential function, to get δm_Z^{fit} (measured m_Z resolution). Finally the average predicted δm_Z is calculated in each $\delta m_Z/m_Z$ bin before and after the correction factor for lepton p_T error is applied (predicted m_Z resolution).

In the closure plot, it is expected to see δm_Z gets closer to δm_Z^{fit} after correction, and the points should stay in a band which is 20% around diagonal line, which is the uncertainty assigned to the resolution in the previous analysis [16]. This closure test is shown in Fig. 5-22 for muons and in Fig. ?? for electrons. A further check has been also performed, looking at the closure test of the predicted four lepton mass resolution compared to the fitted four lepton mass resolution using ggF signal MC samples once the corrections derived using Z events are applied. After applying correction, measured m_{4l} resolution gets closer to the prediction. This closure test is shown for three different final states in Fig.

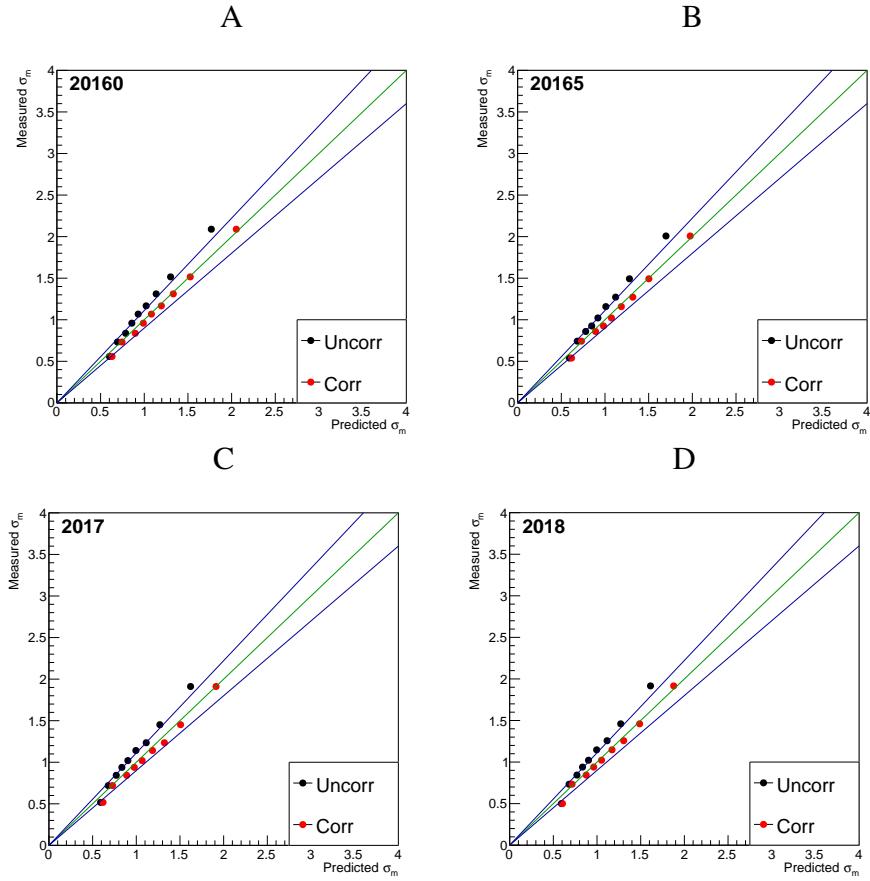


Figure 5-22. Validation of the per-event mass uncertainties from Z events in MC in the dimuon channel using 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right) events. The solid blue line corresponds to a 10% band relative to the solid green 1:1 line.

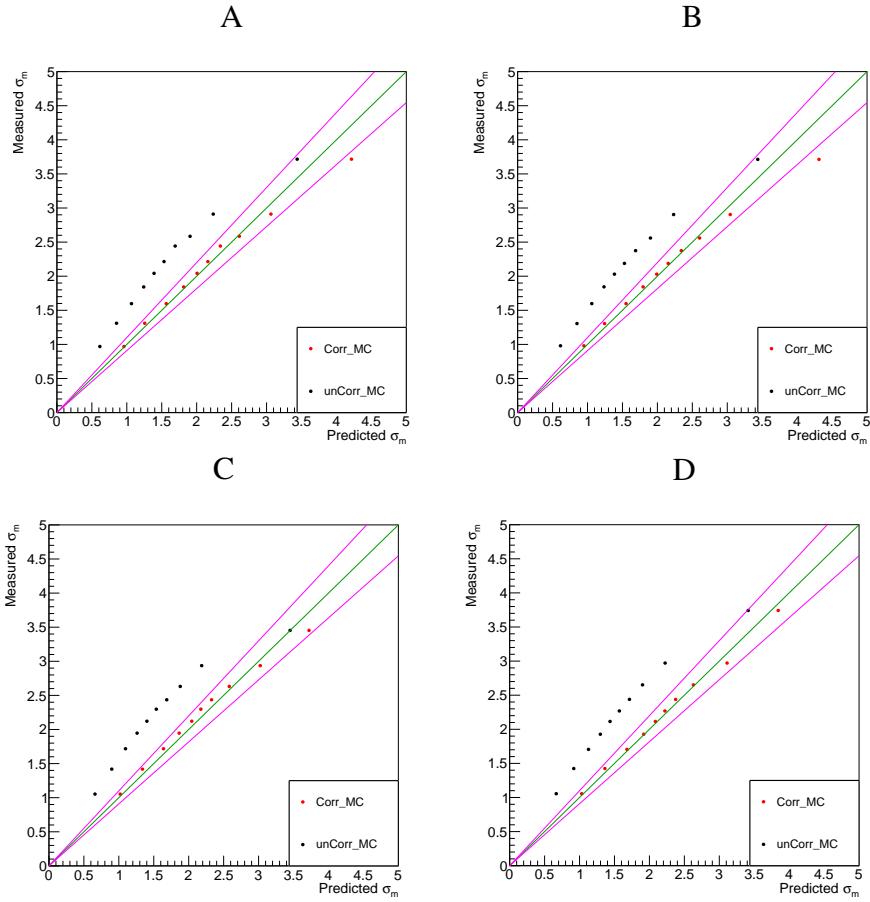


Figure 5-23. Validation of the per-event mass uncertainties from Z events in MC in the dielectron channel using 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right) events. The solid blue line corresponds to a 10% band relative to the solid green 1:1 line.

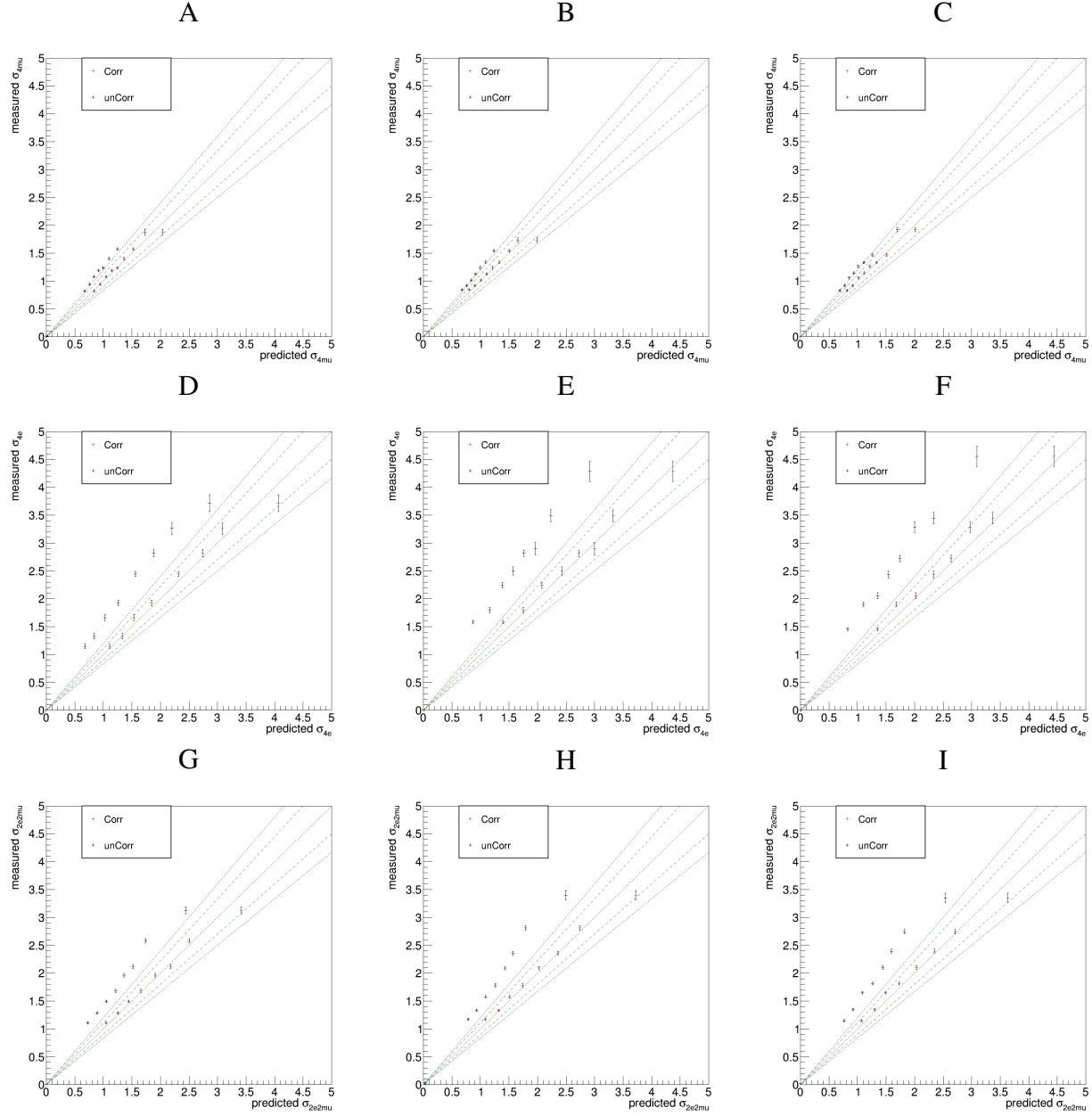


Figure 5-24. Validation of the per-event mass uncertainties from events in $gg \rightarrow H \rightarrow 4\ell$ channel, in MC, for three different final states (4μ on top, $4e$ in the middle and $2e2\mu$ on bottom), for three years (2016 on the left, 2017 in the middle, 2018 on the right). The 20% reference band is also shown.

5.7 Signal Modeling

5.7.1 Signal Normalization

The normalization of the Higgs boson signal is obtained, from simulation, looking at the expected signal yields in the range [105, 140] GeV, for five simulated mass points (120, 124, 125,

126 and 130 GeV). A second order polynomial function is used to extract the dependence of the normalization from m_H . Fits are performed separately for each production mode, for each decay channel and for each year. Examples of the fits can be observed in Fig. 5-25

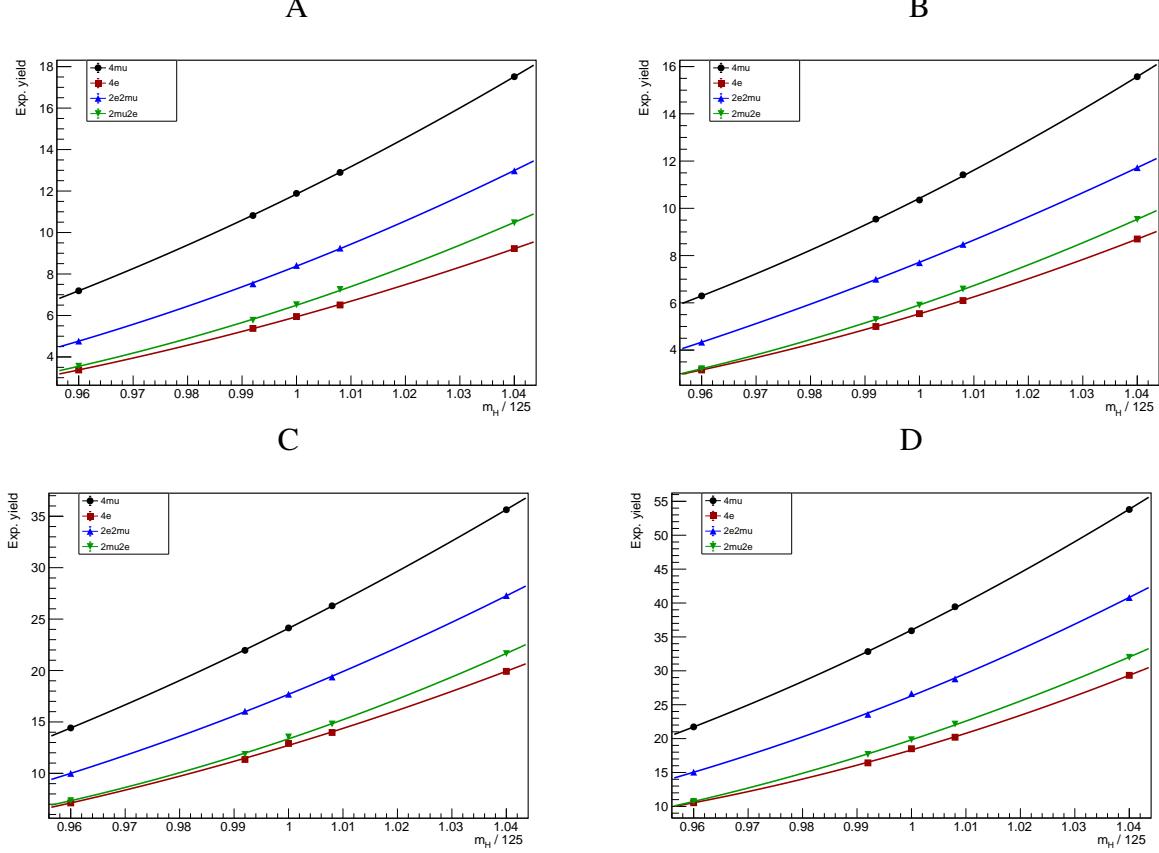


Figure 5-25. Normalization fit for ggH, in different decay channels, as a function of mass, for 2016 on top (pre-VFP on the left, post-VFP on the right), 2017 on bottom left, 2018 on the bottom right.

5.7.2 Parameterizing the Signal Mass Line Shape

The signal lineshape is obtained from the fit of the Higgs boson mass distribution, in the range [105, 140] GeV, using a double-sided Crystal Ball (DSCB) function, which has 6 parameters. Fit parameters are derived as a function of mass, using a first order polynomial:

$$param_{DSCB}^i = a^i + b^i (m_H - 125)$$

First, from the fit of the only 125 GeV sample, the “a” term for each parameter is extracted (“b” term in this case is not taken into account). Then, a second fit is performed: this time, “a” is

fixed to the value found before, while “b” is kept free to float when fitting all five different mass points (120, 124, 125, 126, 130 GeV) simultaneously.

The fit is performed separately, for each production mode, for each decay channel, in each year. To take into account the non-resonant contribution in the case of VH and ttH production modes, the DSCB is convoluted with a Landau function that describes the possibility for a lepton from the Higgs boson decay to be lost or not selected.

5.7.3 Building the 1D pdf

Higgs boson mass measurement is firstly extracted from a one-dimensional likelihood function $\mathcal{L}(m_{4\ell}|m_H)$, where m_H is fixed to the value of 125 GeV. The model and the normalisation used for the signal are described in [5.7.1](#).

As reference, Fig. [5-26](#) shows the four-lepton invariant mass distribution in the signal region ([105-140] GeV) using ggF events, with the DSCB fit for 2018, split in four different final states.

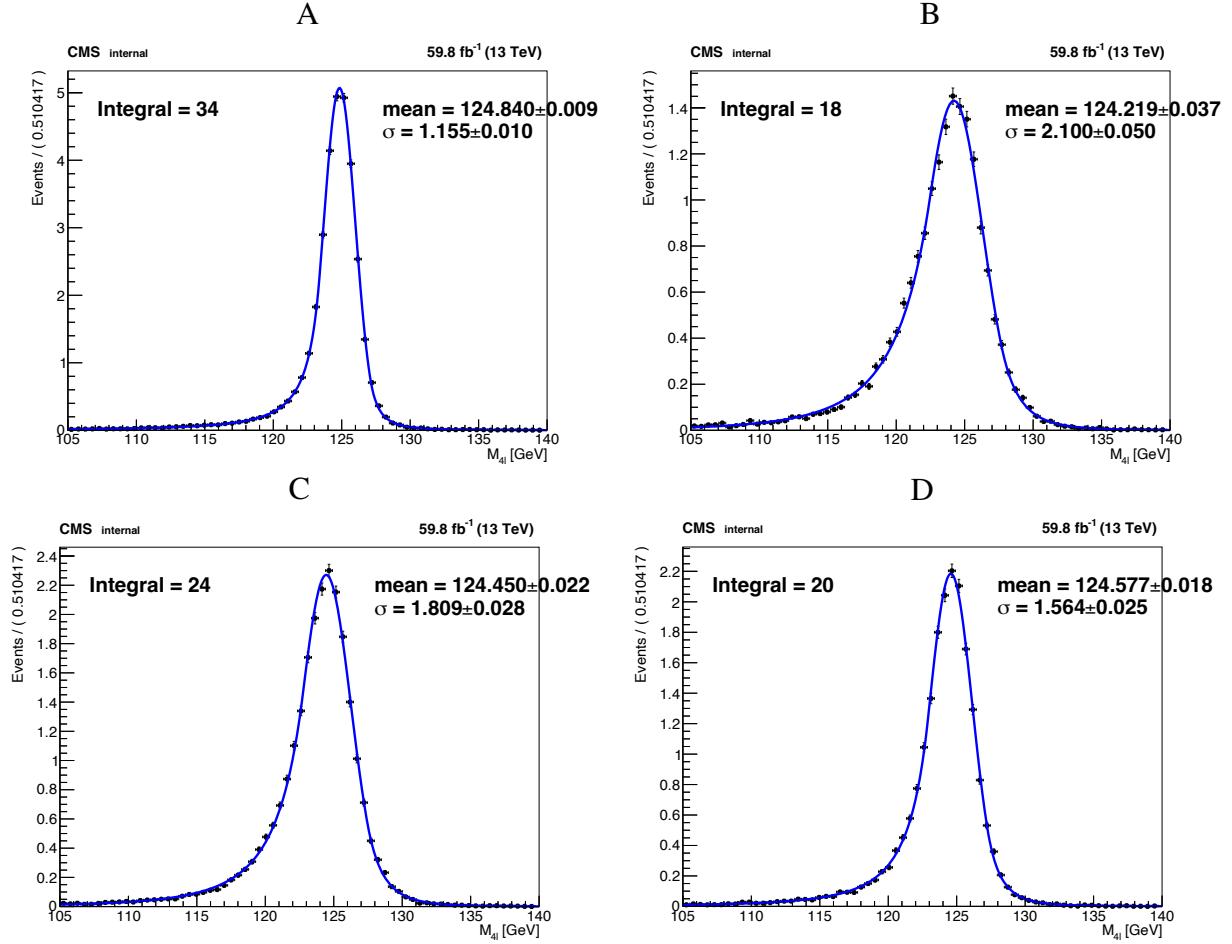


Figure 5-26. Four-lepton invariant mass distribution in the signal region ($[105\text{--}140]\text{ GeV}$) using ggH events, with the DSCB fit for 2018: 4 μ (top left), 4e (top right), 2e2 μ (bottom left), 2 μ 2e (bottom right).

5.8 Systematic Uncertainties

Words.

5.9 Results

Words.

CHAPTER 6

SEARCH FOR LOW-MASS DILEPTON RESONANCES IN THE $H \rightarrow 4\ell$ CHANNEL

6.1 Motivation

As mentioned in Sec. 5.1, even though the Higgs boson has been well studied and *appears* to be consistent with the SM Higgs boson, a single experiment that shows BSM activity (i.e., *any* deviation from SM prediction) is all that is required to defenestrate this idea. For example, it may be the case that the Higgs boson (H) decays into particles other than those found in the SM. This chapter details such an analysis, which follows similar topologies to the one studied in Chapter 5 ($H \rightarrow ZZ^* \rightarrow 4\ell$), specifically $H \rightarrow ZX \rightarrow 4\ell$ and $H \rightarrow XX \rightarrow 4\ell$, where X is a BSM low-mass dilepton resonance.

6.2 Overview

This analysis searches for exotic decays of the Higgs boson (H) using data collected by the CMS experiment (Chapter 4) from the pp collisions produced at the CERN LHC (Chapter 3) at a center-of-mass energy of 13 TeV during the Run 2 period (2016–2018). Two different decay mechanisms are considered: $H \rightarrow ZX$ and $H \rightarrow XX$. Here, X represents a BSM particle that may decay into a pair of OSSF leptons, though only electrons and muons are considered. Because the Z is the SM Z boson, this decay process yields the same four-lepton (4ℓ) final states that were found in the Higgs boson mass measurement analysis (Chapter 5): $4e$, 4μ , $2e2\mu$, and $2\mu2e$. The 4ℓ final state was chosen due to its clean signature and large signal-to-background ratio.

Not only is X a theoretical BSM particle, it is speculated to interact with SM particles via the “dark” sector. Such dark sector particles are described within the theoretical framework of the “hidden Abelian Higgs model” (HAHM) TODO:refs7–11 from EJPC. Within the context of the HAHM, X is the dark photon (Z_D). This dark particle mediates a dark $U(1)_D$ gauge symmetry that is spontaneously broken by a theoretical dark Higgs mechanism.

The search for Z_D begins with a SM decay of $H \rightarrow ZZ$. Then, via the dark sector, one of the Z bosons *mixes* with Z_D , as determined by the kinetic-mixing parameter ε . The Z_D and the other Z then decay into OSSF lepton pairs to yield one of the four different 4ℓ final states. This decay process is shown in Fig. 6-1 (Left). The other mode of Z_D production occurs when a H mixes with a dark Higgs boson (s), as governed by the Higgs-mixing parameter κ . Then, the s decays into two

identical Z_D particles which decay promptly to give one of the 4ℓ final states. This decay process is shown in Fig. 6-1 (Right).

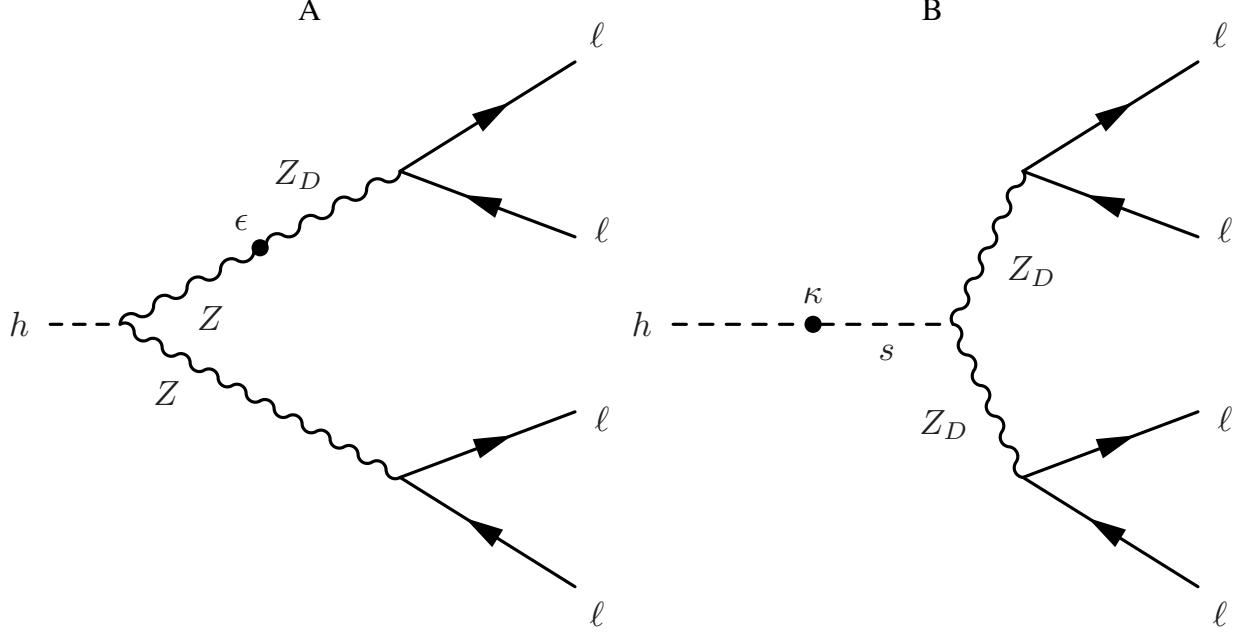


Figure 6-1. Exotic decays of the Higgs boson (h) to the 4ℓ final state within the context of the HAHM. A) Higgs boson decays into two SM Z bosons, one of which kinetically mixes with the dark photon (Z_D). B) Higgs boson mixes with a dark Higgs boson (s) via the Higgs-mixing mechanism, governed by the Higgs-mixing parameter κ . The dark Higgs then decays into two identical dark photons, each of which decays into two OSSF leptons.

This analysis assumes a narrow-width approximation for Z_D decays so that the dark photon is confined to its on-mass shell. Thus, for any given dark photon mass hypothesis, Z_D is assumed to always have that singular mass value. This limitation on the mass of Z_D (m_{Z_D}) in conjunction with the narrow-width nature of the Higgs boson mass resonance ($m_H \approx 125$ GeV) kinematically constrains the search region of the analysis. Specifically, the $H \rightarrow ZZ_D$ ($H \rightarrow Z_D Z_D$) process has the upper limit $m_{Z_D} < m_H - m_Z \approx 35$ GeV ($m_{Z_D} < m_H/2 \approx 62.5$ GeV). Finally, to avoid the ubiquitous but unwanted reconstruction of an intermediate J/ψ ($c\bar{c}$) meson, which has a precisely measured mass of $3.096\,900 \pm 0.000\,006$ GeV [8], the mass range chosen for this analysis is $4.0 < m_{Z_D} < 35.0$ GeV (62.5 GeV).

This low-mass dilepton resonance search has a similar to the aforementioned Higgs boson mass measurement analysis described in

6.3 Data Sets, Simulated Samples, and Triggers

The data sets used for the search for low-mass dilepton mass resonances in SM Higgs boson decays uses the same data sets for all three years (2016–2018) as those found in Ref. ???. A brief explanation of the parameters used to generate the data sets follows next.

The physics processes corresponding to the signal $\text{pp} \rightarrow \text{H} \rightarrow \text{ZZ}_D (\text{Z}_D \text{Z}_D) \rightarrow 4\ell$, where $\ell = (\text{e}, \mu)$, were generated at leading order (LO) using `MADGRAPH5_amc@NLO` 2.2.2 (2.4.2) for 2016 (2017, 2018) samples using the parameters set by the HAHM. On the other hand, `POWHEG v2` was used to simulate the production of SM Higgs bosons via the typical processes (ggH , VBF , VH , and $\text{t}\bar{\text{t}}\text{H}$) and to simulate $\text{q}\bar{\text{q}} \rightarrow \text{ZZ}$ at next-to-leading order (NLO) using perturbative quantum chromodynamics (Refs. [17–20]). The other irreducible background process ggZZ was simulated using `MCFM` 7.0.1 (Ref. [21]) at LO, to which NLO correction factors were applied (Ref. [22]). The simulation of $\text{H} \rightarrow 4\ell$ was carried out by the software `JHUGEN` 7.0.2 (Refs. [23, 24]).

6.4 Event Selection and Reconstruction

6.5 Backgrounds

6.6 Systematic Uncertainties

6.7 Results

An analysis of the m_{Z_2} spectrum is performed to look for any possible low-mass dilepton resonances. In the case of $\text{H} \rightarrow \text{Z}_D \text{Z}_D$, in which both of the daughter particles are identical, then a peak in the m_{Z_2} spectrum is expected at $(m_{\text{Z}_1} + m_{\text{Z}_2})/2$.

A simple counting experiment is performed in many bins across the m_{Z_2} spectrum. Using events selected from the ZZ_D event selection, 353 mass hypotheses m_i are considered for m_{Z_2} . The idea is to scan over the entire m_{Z_2} range (4.20–34.98 GeV) in very fine m_{Z_2} bins, while avoiding the Υbb bound states. To achieve the desired bin width fineness, each subsequent mass hypothesis is increased by 0.5% of its previous value. Thus, the mass hypotheses are given by:

$$m_i = 4.20 \times 1.005^i \text{ GeV}, \text{ where } i = 0, 1, 2, \dots, 129, 202, 203, 204, \dots, 425.$$

The bin width is chosen to be two times the m_{Z_2} resolution. Concretely, the bin width is equal to

0.04 (0.10) $\times m_i$ for the 4μ and $2e2\mu$ ($4e$ and $2\mu2e$) final states.

For each m_i , an overall likelihood model (\mathcal{L}_{m_i}) is defined as:

$$\mathcal{L}_{m_i} = \mathcal{L}_{m_i}^{\text{SR}} \mathcal{L}_{m_i}^{\text{sb}},$$

where $\mathcal{L}_{m_i}^{\text{SR}}$ is the likelihood that the parameters of interest (θ_k) describe the number of events $(n_{m_i, \ell}^{\text{SR}})$ found inside the signal region (SR) for this m_i in a given final state (ℓ), and similarly, $\mathcal{L}_{m_i}^{\text{sb}}$ is the likelihood that the same parameters describe the number of events $(n_{m_i, \ell}^{\text{sb}})$ found inside the sidebands (sb)—i.e., outside the SR—for this m_i in a given final state (ℓ).

Both likelihoods for a given m_i are themselves products of Poisson probabilities¹, which are defined as:

$$\mathcal{L}_{m_i}^{\text{SR}} = \prod_{\ell} \text{Po}\left(n_{m_i, \ell}^{\text{SR}} \middle| \mu n_{s, m_i, \ell} \rho_{s, m_i, \ell} + \mu_H n_{H, m_i, \ell} + \sum_b n_{b, m_i, \ell} \rho_{b, m_i, \ell}\right)$$

and

$$\mathcal{L}_{m_i}^{\text{sb}} = \prod_{\ell} \text{Po}\left(n_{\ell}^{\text{sb}} \middle| \mu_H n_{H, \ell} + \sum_b n_{b, \ell} \rho_{b, \ell}\right),$$

where μ is the signal strength parameter,

n_{ℓ} is the number of

CL_s [25].

¹If the number of expected events (on average) is λ , then the probability to observe x events is given by the Poisson distribution: $\text{Po}(x \mid \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}$

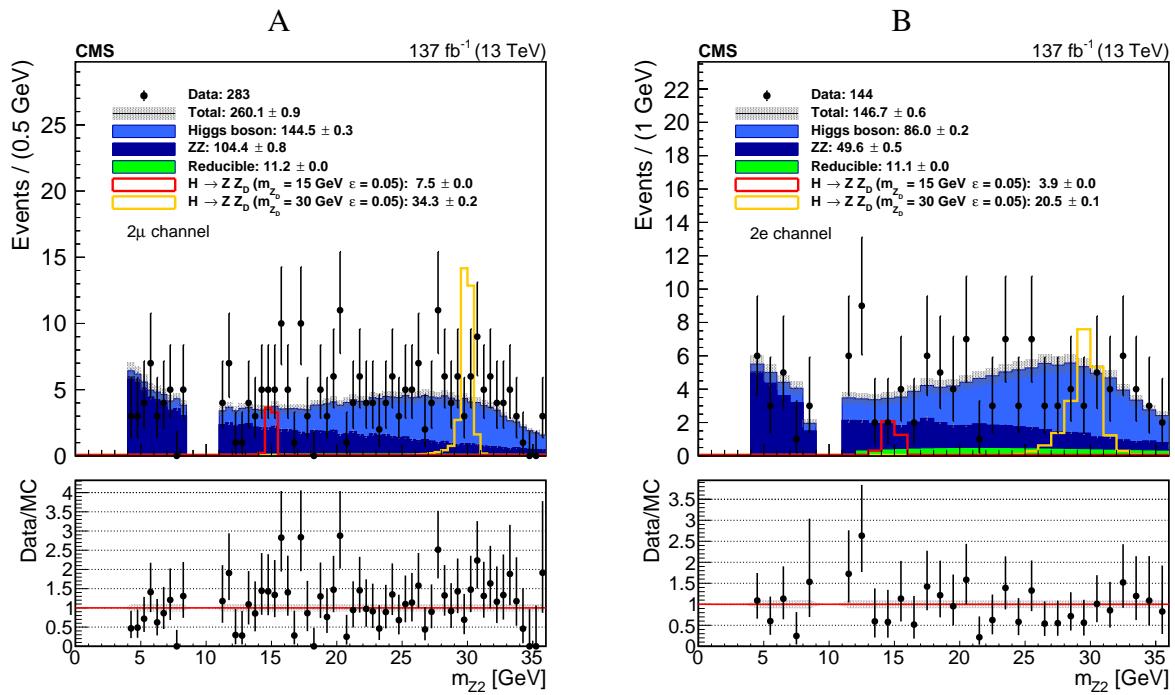
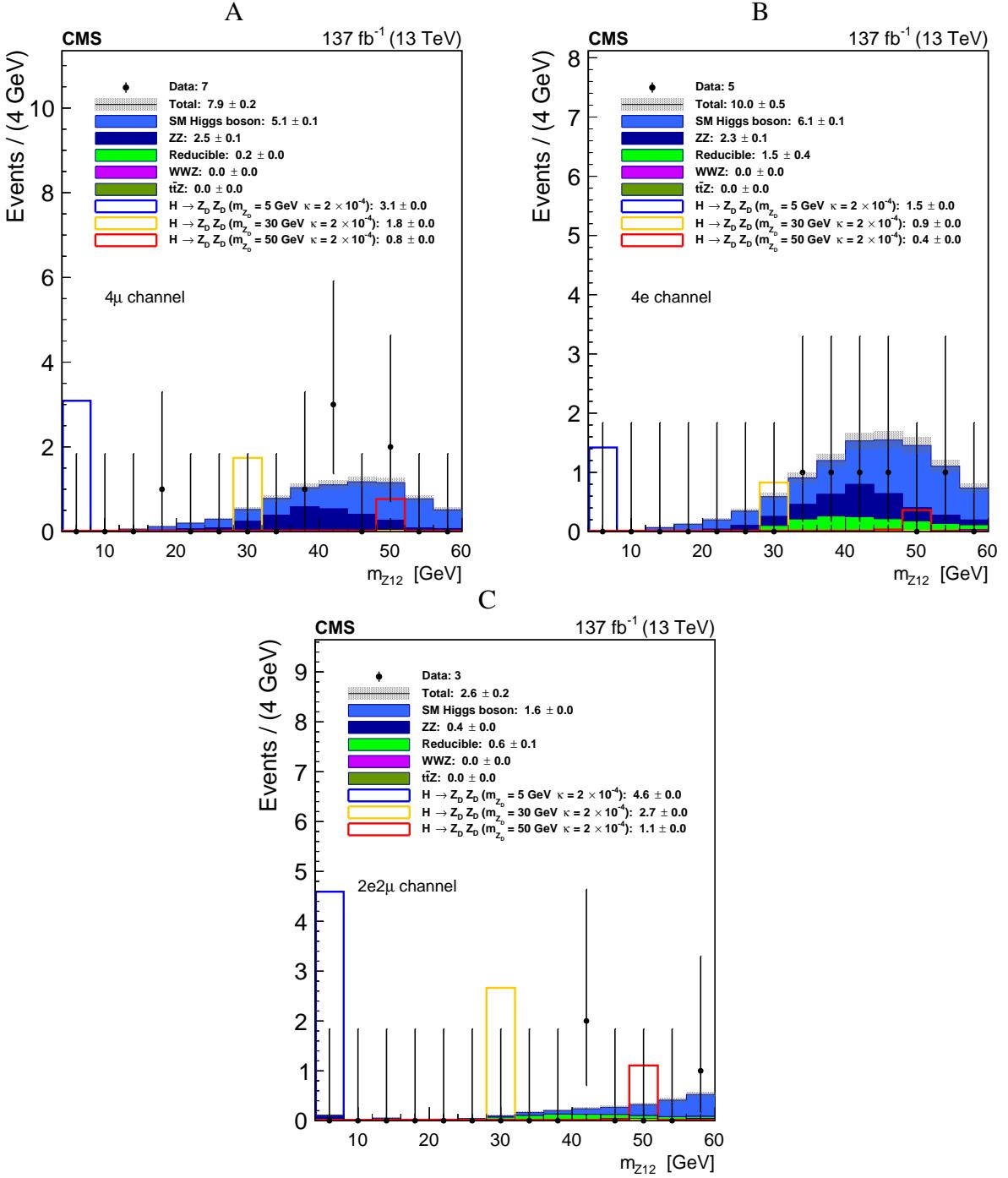


Figure 6-2. testing more



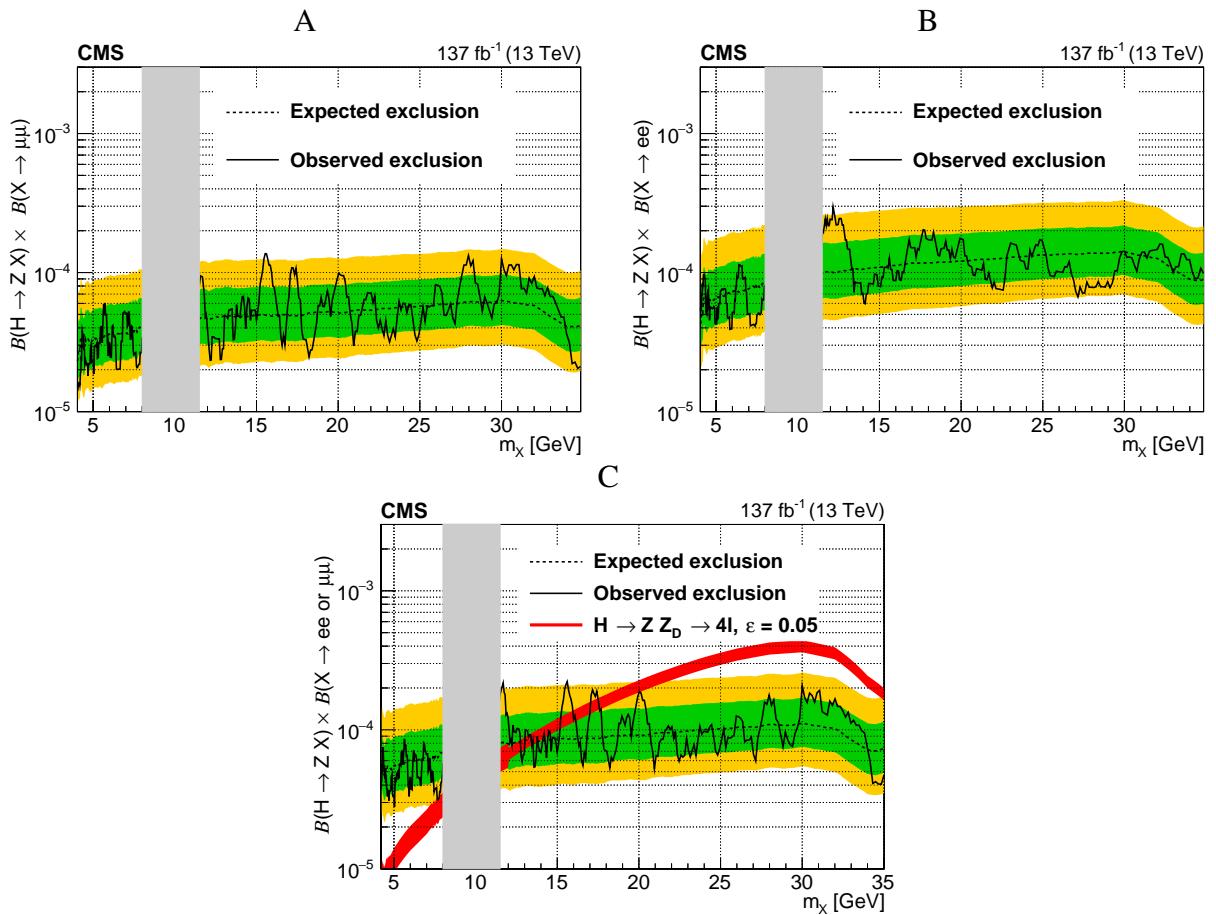
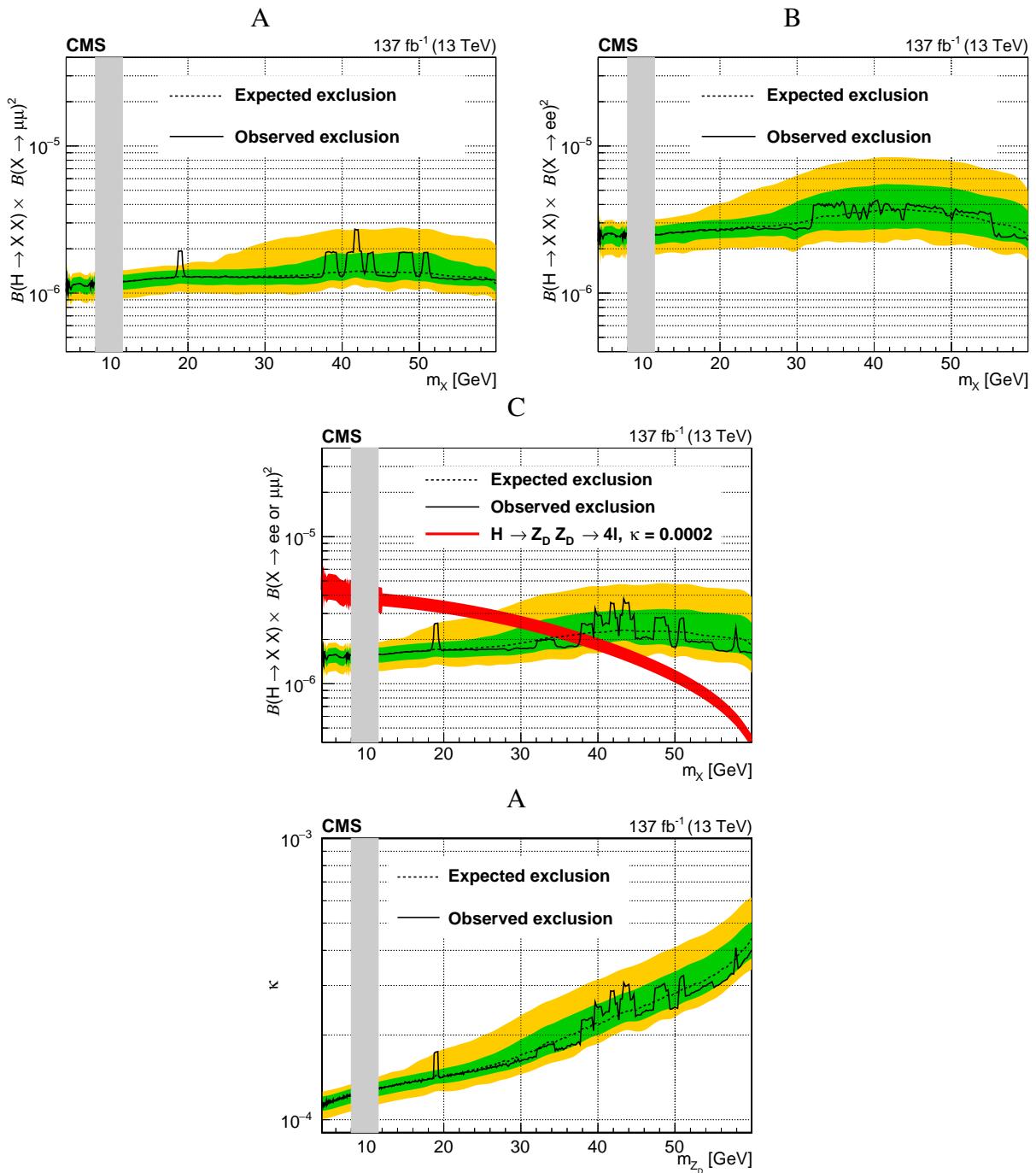


Figure 6-3



CHAPTER 7 CONCLUSION

This dissertation outlines two analyses performed using pp collision data collected by the CMS detector during the LHC Run 2 period (2016–2018), corresponding to an integrated luminosity of 137.1 fb^{-1} . The first analysis details the steps required to measure the mass of the Higgs boson (m_H) in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel, where ($\ell = e, \mu$). The measurement of m_H is expected to be $m_H = 125.38 \pm 0.11 [\pm 0.11 (\text{stat.}) \pm 0.02 (\text{syst.})]$ GeV. The analysis improved upon previous measurements of m_H by constraining muon tracks to a vertex that is compatible with the beamspot (the so-called *vertex+beamspot constraint*). The second analysis considers the decay of a SM Higgs boson to BSM dilepton mass resonances, specifically a dark photon (Z_D), within the context of the Hidden Abelian Higgs Model considering the decay channel $H \rightarrow ZZ_D (Z_D Z_D) \rightarrow 4\ell$. Upper limits on the branching ratios of $\mathcal{B}(H \rightarrow ZZ_D)$, $\mathcal{B}(H \rightarrow Z_D Z_D)$, and $\mathcal{B}(Z_D \rightarrow e^+ e^- \text{ or } \mu^+ \mu^-)$ are set at the 95% confidence level as well as the Finally, upper limits on the Higgs-mixing parameter (κ) are set at the 95% confidence level. No significant deviations from the predictions made by the SM are observed.

REFERENCES

- [1] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012) 30.
- [2] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer et al., *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *Journal of High Energy Physics* **2013** (2013) 81.
- [3] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A.A. Abdelalim et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012) 1.
- [4] ATLAS COLLABORATION AND CMS collaboration, *Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments*, *Phys. Rev. Lett.* **114** (2015) 191803.
- [5] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, T. Bergauer, M. Dragicevic et al., *A measurement of the Higgs boson mass in the diphoton decay channel*, *Physics Letters B* **805** (2020) 135425.
- [6] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. Abidi et al., *Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, *Physics Letters B* **784** (2018) 345.
- [7] A. Andreassen, W. Frost and M.D. Schwartz, *Scale-invariant instantons and the complete lifetime of the standard model*, *Phys. Rev. D* **97** (2018) 056006.
- [8] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *Progress of Theoretical and Experimental Physics* **2020** (2020) 083C01.
- [9] CDF collaboration, *High-precision measurement of the W boson mass with the CDF II detector*, *Science* **376** (2022) 170.
- [10] V.V. Barinov, B.T. Cleveland, S.N. Danshin, H. Ejiri, S.R. Elliott, D. Frekers et al., *Results from the Baksan Experiment on sterile transitions (BEST)*, *Phys. Rev. Lett.* **128** (2022) 232501.
- [11] CMS collaboration, *The CMS experiment at the CERN LHC*, *J. Inst.* **3** (2008) S08004.
- [12] CMS collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, *J. Inst.* **10** (2015) P06005.
- [13] CMS collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, *Journal of Instrumentation* **9** (2014) P10009.
- [14] W. Adam et al., *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, *Journal of Physics G: Nuclear and Particle Physics* **31** (2005) N9.
- [15] CMS collaboration, *Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV*, *J. Inst.* **7** (2012) P10002.

- [16] CMS collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **11** (2017) 047 [[1706.09936](#)].
- [17] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *J. High Energy Phys.* **2007** (2007) 070.
- [18] E. Bagnaschi, G. Degrassi, P. Slavich and A. Vicini, *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*, *J. High Energ. Phys.* **2012** (2012) 88.
- [19] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *J. High Energ. Phys.* **2010** (2010) 43.
- [20] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, *J. High Energy Phys.* **2004** (2004) 040.
- [21] J. Campbell and T. Neumann, *Precision phenomenology with MCFM*, *J. High Energ. Phys.* **2019** (2019) 34.
- [22] M. Grazzini, S. Kallweit and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, *Eur. Phys. J. C* **78** (2018) 537.
- [23] Y. Gao, A.V. Gritsan, Z. Guo, K. Melnikov, M. Schulze and N.V. Tran, *Spin determination of single-produced resonances at hadron colliders*, *Phys. Rev. D* **81** (2010) 075022.
- [24] S. Bolognesi, Y. Gao, A.V. Gritsan, K. Melnikov, M. Schulze, N.V. Tran et al., *Spin and parity of a single-produced resonance at the LHC*, *Phys. Rev. D* **86** (2012) 095031.
- [25] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554.

BIOGRAPHICAL SKETCH

Jake Rosenzweig had the best childhood anyone could ask for, growing up in Jacksonville, FL: enjoying video games with excellent friends, playing football on the beach, and having plenty of opportunity to make mistakes. He graduated from the University of Florida in 2011 with a B.S. in chemistry, while maintaining his sanity by getting minors in education and Latin. He enjoys building things from scrap, weightlifting, hiking in the Coloradoan mountains, gardening, silence, and—most of all—receiving the beleaguered stare from his wife after telling her a *particularly* bad dad joke.